

1988

Testing academic listening skills

Mei Kao Hsueh
Iowa State University

Follow this and additional works at: <https://lib.dr.iastate.edu/rtd>



Part of the [Bilingual, Multilingual, and Multicultural Education Commons](#)

Recommended Citation

Hsueh, Mei Kao, "Testing academic listening skills" (1988). *Retrospective Theses and Dissertations*. 281.
<https://lib.dr.iastate.edu/rtd/281>

This Thesis is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

Testing academic listening skills

by

Mei Kao Hsueh

A Paper Submitted to the
Graduate Faculty in Partial Fulfillment of the
Requirements for the Degree of
MASTER OF ARTS

Department: English
Major: TESL

Signature redacted for privacy

Iowa State University
Ames, Iowa

1988

TABLE OF CONTENTS

	Page
PART I. INTRODUCTION	1
The Specific Problem	1
The Listening Comprehension Test of the present English-As-Second-Language Placement Test (ESLPT)	2
A Closer Look at the Current Trends of ESL Teaching and Testing	4
Communicative language approach	4
English for specific purpose (ESP)	11
Toward authenticity	13
Listening comprehension	15
Testing listening comprehension	19
PART II. THE NEW ESLPT LISTENING COMPREHENSION TEST DEVELOPMENT	23
Stage 1: Planning Stage	23
Stage 2: Pre-pilot Stage	27
Stage 3: Pilot Stage	29
Stage 4: Validation Stage - Concurrent Testing	39
PART III. CONCLUSIONS AND SUGGESTIONS FOR FURTHER REVISION	52
BIBLIOGRAPHY	55
APPENDIX A: SCRIPT OF THE NEW ESLPT	59
APPENDIX B: TESTBOOK OF PILOT VERSION	89
APPENDIX C: TESTBOOK OF FINAL VERSION	97
APPENDIX D: STEPS OF ADMINISTERING THE NEW ESLPT LISTENING COMPREHENSION TEST	104

Part I: Introduction

The Specific Problem

Since the early 1980's, the present English-As-a-Second-Language Placement Test (ESLPT), a post-admissions test, has been used to determine whether an incoming international student (either graduate or undergraduate) needs more work in English to meet ISU standards (these students begin academic work whether or not they pass the ESLPT). Students whose native language is not English are required to take the ESLPT except for those who have taken it as an undergraduate. Hence, most of them take the test a few days before their first semester at the university.

The ESLPT consists of four parts - Part 1: a composition; Part 2: a listening comprehension test; Part 3: a reading comprehension test; and Part 4: a vocabulary test. All but Part 1 are multiple-choice format tests. Based on how the students do on the subtests they are placed in one or more courses - English 100B, C, D, or E. English 100B is an intermediate grammar review and composition course, 100C an advanced composition course for undergraduates, 100D an advanced composition course for graduate students, and 100E supervised independent study which has two sections - section 1: listening comprehension, and section 3: reading comprehension. Those students who are assigned to take 100E courses are required to take them during their first semester of enrollment to ensure that they are better prepared to do academic work.

The trend of English as a second/foreign language teaching and testing theory, as well as practice, has moved in recent

years from a focus on linguistic proficiency to one on communicative competence, shifting from an emphasis on language form to language use, using authentic language, language in the real world, in teaching and testing materials is considered necessary under the communicative language teaching and testing approach. In a related development, English for Specific Purpose (ESP) has emerged as a particular subdivision of Teaching English to Speakers of Other Languages (TESOL). In ESP learners' needs and special interests have been considered by the contribution of a new developments in educational psychology which emphasize the central importance of the learners and their attitudes toward learning (Hutchinson 1987). A subdivision of ESP, English for Academic Purpose (EAP), is based on ESL learner's needs of using English to pursue academic work in institutions where instructions are given in English. The focus of this paper is on testing listening proficiency of ISU post-admissions international students, in other words, testing listening skills related to academic demands, is the essential, specific issue in the paper.

The Present ESLPT Listening Comprehension Test

The present ESLPT listening comprehension test is a tape-recorded test. It consists of 3 parts - Part A, B, and C, comprising 35 multiple-choice test items. In Part A, the students are asked to listen to 8 brief statements, each statement only spoken once. After they have listened to each statement they are asked to choose one statement (out of four) that is closest to the meaning to what they have heard. An example is:

The students hear...

The Wagners seen to take their large family in stride.

The students read...

- A. Mr. Wagner comes from a large family.
- B. Mr. and Mrs. Wagner are the parents of many children.
- C. The Wagners dislike small families.

Part B of the test has 17 short conversations between two people. When each conversation ends, the students are asked some questions by about the conversation by a third voice, and need to choose the answer from the choices given. An example is:

(Man) "I have never understood how to use the card catalogue, and I need to get this information today."

(Woman) "Why don't you pick up one of the guidebooks at the front desk? It will help you, I'm sure."

(3rd voice) "Where did this conversation most likely occur?"

The students will read:

- A. At a museum.
- B. In a store.
- C. In a library.
- D. In a classroom.

In Part C of the test, the students are asked to listen to three short lectures. After they have heard each lecture, they are asked some questions about what was said. All together there are 10 test items in this part.

Recently, the listening comprehension part of the ESLPT at ISU has gone through careful scrutiny in light of the current ESL teaching and testing theory and practice. As a result, a new version of the ESLPT Listening Comprehension test was developed during Fall 1987. It was pilot-tested in December, and problematical questions were subsequently revised based on the

results. In January 1988, a trial test was given concurrently with the present ESLPT listening comprehension test to ISU spring semester incoming international students. A literature review on the current language teaching and testing theory and practice of listening comprehension, along with the procedures and results of the pilot and trial testing, as well as a description of the new ESLPT Listening Comprehension test will be reported respectively in the latter parts of this paper.

A Closer Look at the Current Trends of ESL Teaching and Testing Communicative Language Approach

"An extended definition, or theory, of communicative competence has important implications for understanding a communicative approach to language teaching, ... (Brown, 1987, p.200)". Canale and Swain (1980) have examined various theories of communicative competence that have been proposed, and their definition is now the reference point for virtually all discussions of communicative competence (Brown 1987). In proposing their theoretical framework they have made several general assumptions about the nature of communication and of a theory of communicative competence. They say:

Following Morrow (1977), we understand communication to be based in sociocultural, interpersonal interaction, to involve unpredictability and creativity, to take place in a discourse and sociocultural context, to be purposive behavior, to be purposive behavior, to be carried out under performance constraints, to involve use of authentic (as opposed to textbook-contrived) language, and to be judged as successful or not on the behavioral outcomes. We assume with Candlin (1978) that the relationship between a proposition (or the literal meaning of an utterance) and its social meaning is variable across different socioculture and discourse contexts, and that communication involves the continuous evaluation and negotiation of social meaning on the part of the participants (Canale and Swain, 1980, p.29).

They also agree with Palmer (1978) that genuine communication involves the "reduction of uncertainty" on behalf of the participants, and assume that communication involves verbal and nonverbal symbols, oral and written modes, and production and comprehension skills. As to the theory of communicative competence, they assume that a theory of communicative competence interacts with a theory of a human action and with other systems of human knowledge (e.g. world knowledge). In Canale and Swain's (1980) definition, and later in Canale's (1983) definition, four different components make up the construct of communicative competence (Brown 1987). These components are:

1. Grammatical competence. This type of competence includes knowledge of lexical items and of rules of morphology, syntax, sentence-grammar semantics, and phonology.
2. Discourse competence. This is the ability to connect sentences in stretches of discourse and to form a meaningful whole out of a series of utterances.
3. Sociolinguistic competence. This is the knowledge of two sets of rules, sociocultural rules of use and rules of discourse.
4. Strategic competence. According to Canale and Swain (1980), this involves "verbal and nonverbal communication strategies that may be called into action to compensate the breakdowns in communication due to performance variables or to insufficient competence (p. 31)".

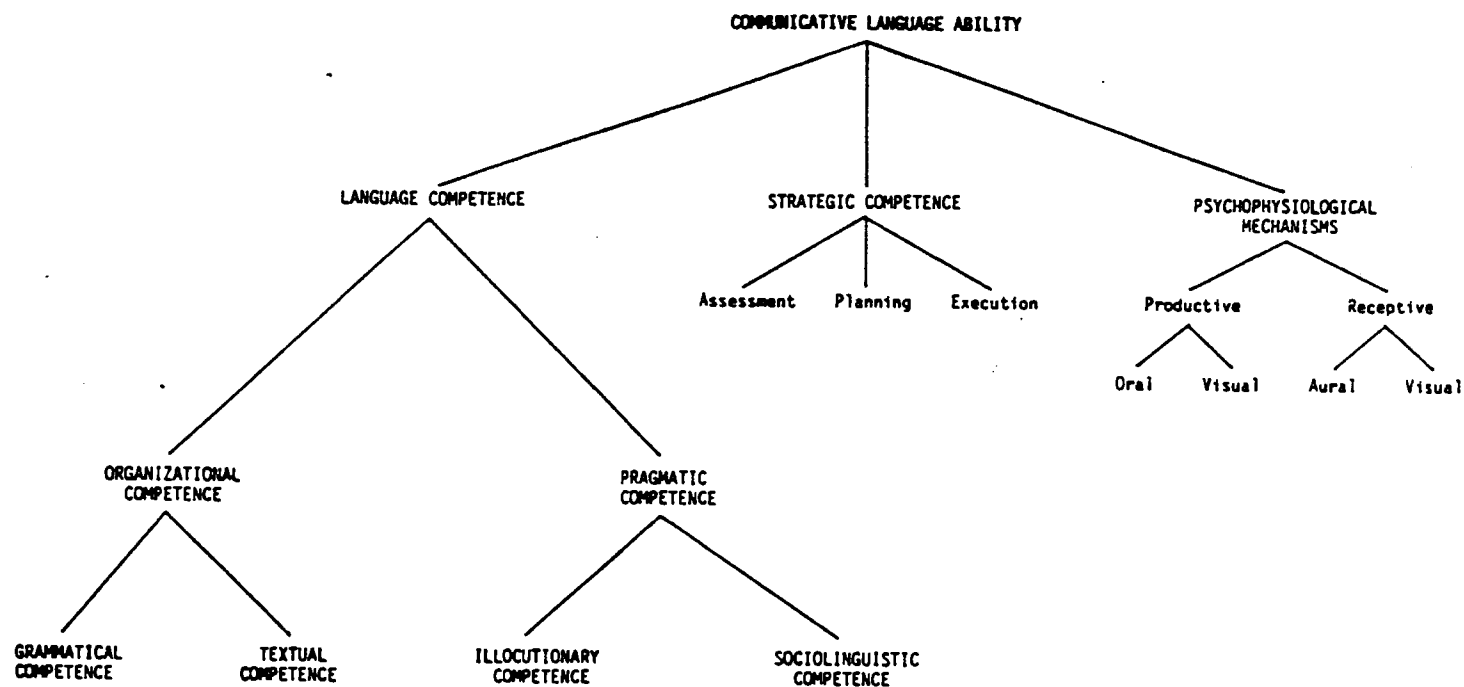
Even though Canale and Swain's definition is often referred to when discussing communicative competence, in Brown's (1987) opinion, in the near future, another model developed by Bachman

(1979) will be widely adopted by researchers, since his model stresses the importance of strategic competence, which would appear to have more descriptive adequacy than Canale and Swain's. In Bachman's model (1988), there are three elements of "communicative language ability" (CLA) - language competence, strategic competence, and psychomotor mechanisms required in executing these competencies in language use (see Figure 1). Language competence consists of organizational and pragmatic competence, and organizational competence comprises the knowledge of grammatical and textual rules which corresponds to Canale and Swain's "grammatical competence" and "discourse competence". As to pragmatic competence, it includes illocutionary competence (the knowledge of how to perform language functions) and sociolinguistic competence. Bachman further defines strategic competence as consisting of three components: assessment, planning, and execution. His strategic competence is the ability to use all the elements of language competencies available, including psychomotor mechanisms (sensory and motor mechanisms), in achieving, realizing the communicative goal.

In addition, Oller (1987) in his effort to define communicative competence points out that, in spite of the popularity of the dyadic definition of communication, "communication is an interaction between n persons where n must be equal or greater than two (p.107)", a more general definition of communication on which to base a theory of communicative competence seemed to be more appropriate. Oller's communicative competence (or general intelligence) model shows the

Figure 1 A framework for Describing Communicative Language Proficiency (Bachman 1988).

A FRAMEWORK FOR DESCRIBING COMMUNICATIVE LANGUAGE ABILITY



interrelation of various mental constructs as components of communicative competence (see Figure 2). The theoretical framework of communicative competence proposed by Oller suggests that communicative competence "is a dynamic state that obtains within a hierarchy of semiotic systems to the extent that those systems are well equilibrated with reference to each other and to externality (p.118)".

Candlin (1987) argues against a traditional (descriptive) view of communicative competence, which leads to conformity to norms, and sympathizes with an alternative (explanatory) view which understands communicative competence as "capacity": "the ability to create meanings by exploring the potential inherent in any language for continual modification in response to change, negotiating the value of convention rather than conforming to established principle (p.39-40)". To Candlin the Hallidayan view of language, which focuses on the interdependence of text, ideation, and interpersonality (Halliday 1979), is seen to have significant value (see Figure 3).

In conclusion, these theorists, and many others, all contribute to our understanding of what communicative competence is. It is the ability, or "general intelligence" as termed by Oller, or "capacity" as termed by Candlin, which allows a person to activate his/her language and world knowledge, as well as various skills (e.g. psychomotor skills) to communicate effectively, produce and receive meaningful signals, whether verbal or nonverbal, to fulfill specific communication purposes. Teaching and testing communicative competence is the ultimate objective in many ESL programs today. The

Figure 2

A slightly elaborated model of the interrelation of various mental constructs as components of general intelligence, alias communicative competence.

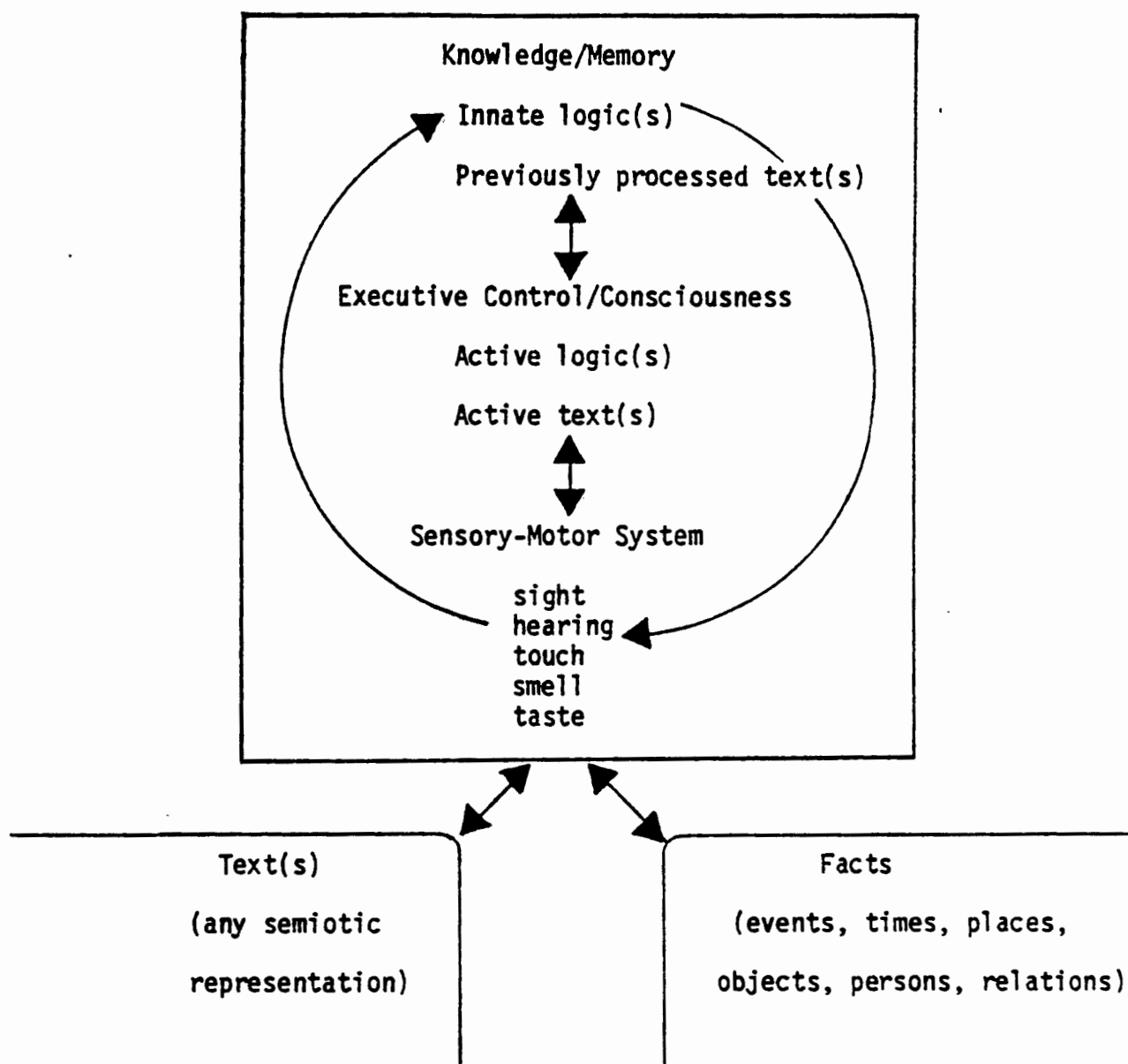
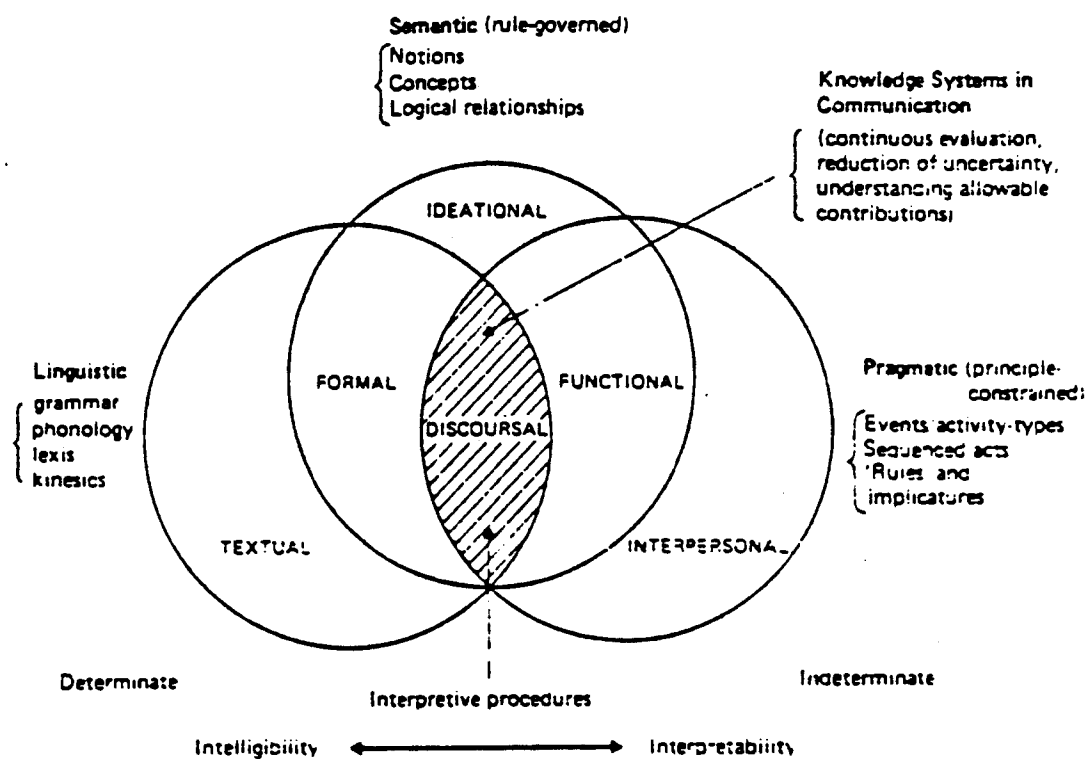


Figure 3

A Hallidayan Model of Communicative Competence

From Language as Social Semiotic by M. A. K. Halliday, 1979,
London: Arnold.



Problems facing Interpretability

- | | | |
|------------------------------|---|------------------------------------------------|
| 1. Textual (Linguistic) | } | 4. Discoursal "capacity"
(Psycholinguistic) |
| 2. Ideational (Semantic) | | |
| 3. Interpersonal (Pragmatic) | | |

theoretical frameworks of communicative competence mentioned above could serve as a base for communicative language test writers to justify or evaluate the content, procedures, testing techniques, and scoring method developed for their tests.

English for Specific Purpose (ESP)

For the purpose of designing the new version of ISU ESLPT Listening Comprehension Test, special attention should be given to English used for academic purposes (EAP). Since EAP is known as a subdivision of ESP, a discussion on ESP for a deeper understanding of the development, definition, and issues related to it seems necessary. According to Hutchinson and Waters (1987), ESP is a phenomenon that has grown out of a number of converging trends: First, expansion in scientific, technical, and economic activity on an international scale, has created a demand for an international language. Second, the revolution in linguistics has shifted the focus on grammar (language rules) to the discovery of ways in which language is actually used in real communication. Third, new developments in educational psychology have centered on learners' different needs and interests. All of these have led to a need for increased specialization in language teaching and testing. Over the years, ESP has gone through several developmental periods from the early "register analysis", "rhetorical analysis", "target situation analysis" (e.g. J. Munby in Communicative Syllabus Design [1978]), to the thinking processes that underlie language use (skills and strategies).

What is ESP exactly? To provide a deeper understanding of ESP, the question could be answered from several angles. First, how does ESP differ from general purpose English (GPE)? Widdowson

(1983) in making a distinction between ESP and GPE, points out that in ESP, "purpose" refers to the eventual practical use to which the language will be put in achieving occupational and academic aims (EOP and EAP). According to him, "ESP is essentially a training concept", after one has established as precisely as possible what learners need the language for." But, in GPE, purpose has to be conceived of in educational terms, as a formulation of objectives which will achieve a potential for later practical use. In other words, purpose is a descriptive term in ESP, while in GPE, it is a theoretical term.

Secondly, Hutchinson and Waters (1987) point out that ESP is not dealing with "specialized varieties" of English (product of register analysis). Similarly, Mackay and Mountford (1978) point out that "ESP implies a special aim, this aim may decide the area of language required, skills needed and the range of function to which language is to be put (p.4)", thus ESP is dealing with specialized aim not special language.

Thirdly, how do we define the varieties within ESP (e.g. EST, EAP, EOP, etc.), given the fact that there are so many divisions? Basically, two kinds of devices have been developed for defining the varieties: purpose - register analysis, and needs analysis (Widdowson 1983). Yet Widdowson also pointed out that "in both cases the analysis is reductionist and divisive (p.10)", and proposes a comprehensive theoretical view of ESP. It is suggested that ESP should be arranged along a scale of specificity with training at one end and education at the other, and EAP would appear close to the end of education. Skehan brings

up a similar concern when discussing the issues in ESP testing (1984), questioning the communicative needs analysis approach that Munby (1978) took when developing the language skills taxonomy which has been influential in the production of some new ESP tests (e.g. ELTS, the British Council test [Carroll 1980]). Instead, Skehan supports the notion of unitary competence advocated by Oller (1976, 1983), which appears to have much in common with Widdowson's comprehensive theoretical view of ESP. Furthermore, when considering the various skills needed for some ESP situations, Skehan questions the possibility of a criterion-referenced approach for tests of ESP, pointing out that though criterion referencing is an attractive ideal, it is difficult to achieve in practice. He suggests that ESP testing should either use the concept of criterion-related validity or try to develop appropriate norm-referenced tests, after all "how specific is specific (Skehan 1984, p.213)"?

Toward Authenticity

Since the communicative language and ESP approaches involve the ability to use language to communicate in real life situations for real purposes, an emphasis on authenticity is inevitable. However, as Spolsky (1985) points out, the criterion of authenticity raises important pragmatic and ethical questions. Shohamy and Reves (1985) have surveyed the development of language tests toward authenticity. After analyzing the advantages and shortcomings of indirect vs. direct (authentic) tests, they point out that indirect tests are more efficient and suitable for psychometric analysis. Authentic tests, on the other hand, based on real interactions, and aiming at capturing real

language and its variations in use have more face validity in terms of authenticity. Nevertheless, there is a lack of attention given to the psychometric properties of authentic tests, and it is "a naive belief that so-called authentic tests are really authentic (Shohamy and Reves 1985, p.51)"; the language obtained on the tests is no more than "authentic test language" which is not the same as real-life language (Stevenson 1985, Spolsky 1985). Through the analysis of the differences between authentic test language and authentic real-life language, Shohamy and Reves identified five factors considered responsible for reducing the "authentic" language into "authentic test language". These factors are: the goal of interaction, the participant, the test setting, the topic, and the time of test. Spolsky (1985) has also looked into the issues of authentic tests, noticing some related problems and suggesting that some of the difficulty can be overcome by authentic-seeming task, and long, patient, and sympathetic observation by testers who are more like observers. Seliger (1985) after examining some ways in which meaning is extracted in authentic language context, points out that by the very act of testing, authenticity is destroyed, yet also suggests some implications for authentic language testing, mainly that, "language testing would be to develop tests that focus not on the product, but on the successful implementation of the processes (Seliger 1985, p.12)". Knowing the pragmatic and ethical questions raised on authentic language tests, and acknowledging the fact that "testing" itself is a stumbling block of authentic language testing. Though designing an authentic language test might seem

to be an impossible mission, Selinger's suggestion has at least pointed out a direction for language testers.

Note-taking is often present in the academic environment, and when recall of information is involved many students believe in note-taking (Canke 1981). The results of studies done on the value of note-taking have been inconsistent, some find it beneficial but some do not, as Peters noted (1972), in terms of subsequent test outcomes. Some studies indicated that the effect of note-taking is conditional, depending on factors such as intelligence, verbal ability, and subject's memory span (Peck 1983; Vesta and Gray 1973). Despite the fact that studies dealing with the effect of note-taking on recall offer mixed support for the value of this activity, since note-taking is an activity that students often engage in while listening to information or lecture presented to them in academic environment. Therefore, in EAP listening comprehension tests, for striving toward authenticity, allowing students to take notes is justifiable.

Listening Comprehension

"Listening comprehension is a complex, problem-solving skill (Wipf 1984, p.345)". In recent years, much research has been conducted on the various skills involved in listening, the central place of listening comprehension in initial second language acquisition, and its influence upon the development of the other language skills (Douglas 1988). Dunkel (1986) in her discussion of the theoretical principles and pedagogical considerations of developing listening fluency in a second language has noted that "listening comprehension should be the focal methodology in foreign/second language instruction,

particular at the initial stages of language study (p.99)".

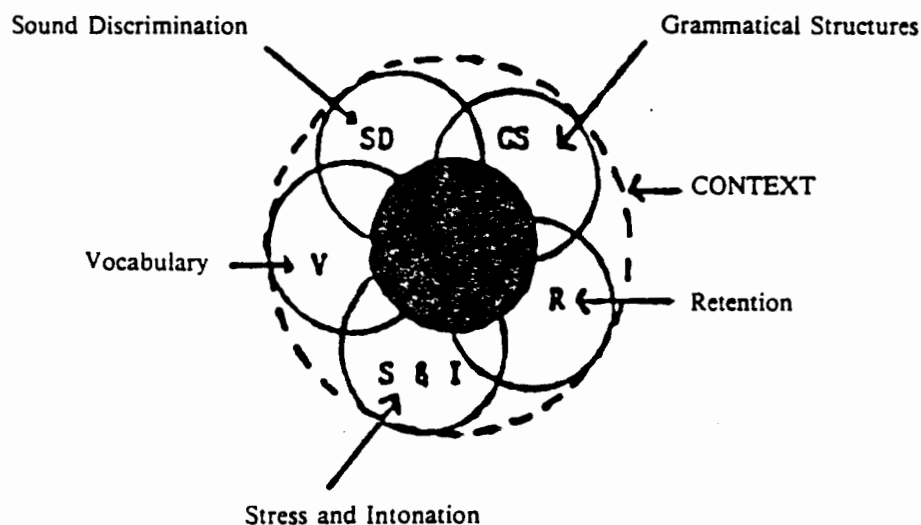
Many others have stated the same idea, acknowledging the importance of listening skills in any communicative situation, and advocating the development of materials for purposes of teaching and testing listening comprehension (Wilkes 1981, Burbidge 1984, Benson 1980, Postovsky 1974, Davies 1980).

Before we turn to a detailed discussion of listening comprehension, what can be said about listening in the real world? According to Burbidge (1984) and Brown (1978), typical spoken language is not very well organized, seldom consists of complete sentences, is full of hesitations, interruptions and repetitions, and often does not contain a lot of information. In real listening, there is usually a lot of visual aids, which helps the listener to comprehend the message. Furthermore, in real world listening people have expectations about how the conversation might go, based on their knowledge of the world and are constantly making predictions about what might be said next. In addition, listeners have a lot of contextual information about the speaker and the situation they are in. And very often they have purposes for listening, seeking for specific pieces of information. Knowing what is involved in real listening will certainly help our understanding about what the components of listening comprehension are.

According to Wilkes (1981), Benson and Hjeltn (1980) agree with Oller (1974) that "listening comprehension is a highly integrative skill which demands conceptualization of the phonological, grammatical and lexical data into an internal

expectancy grammar (p.88)". Rivers (1968), discusses two levels of listening comprehension: (1) the recognition level (the identification of words and phrases, structural interrelationships, time sequences, logical and modifying terms, and others); and (2) the selection level (seizing on those elements of the speaker's language which seem to express the purposes of the speaker). Wipf (1984) through a graphic representation has tried to demonstrate the components of listening comprehension (see Figure 4):

Figure 4: Components of Listening Comprehension



(Figure 4: The dark space in the center, where all the components are superimposed, represents listening comprehension; and the size of the darkened area determines the degree of mastery; and the enveloping circle represents the context in which listening comprehension takes place).

Furthermore, three related levels of discourse processing involved in listening have been identified by Richards (1985).

These processes are: propositional identification (propositions are the basic units of meaning involved in comprehension), interpretation of illocutionary force, and activation of real-world knowledge. Richards also points out that the act of speaking imposes a particular form on utterances, and this considerably affects how messages are understood (medium factors). Based on his analysis, he has come up with a taxonomy of listening skills, including 33 micro-skills related to conversational listening, and 18 relevant to academic listening. Powers (1985), on the other hand, on the basis of a literature review, has come to the conclusion that "there is no comprehensive theory of listening behaviour (p.3)". Despite the lack of a simple, completely acceptable definition of listening, after his literature review, Powers has summarized some factors thought to affect listening comprehension. These factors relate to the stimulus material (i.e., mode of presentation), to the speaker (i.e., the manner of delivery), and the context (i.e., whether or not notetaking is allowed) or other situational variables. The expositions of the researchers mentioned above contribute to our insights into various aspects of listening comprehension. There is no doubt that listening comprehension is a complex, integrative skill. It demands the knowledge of grammatical structures, phonology, and morphology, and more important the ability to apply the knowledge into various listening situations for a proper interpretation of an utterance or a series of utterances, and as a result, lead to a fulfillment of specific listening purposes.

Testing Listening Comprehension

The fact that listening comprehension has been taking on growing importance in foreign/second language learning and teaching has brought to testing specialists an obligation to assist in the development of methods for valid and reliable measurement of listening comprehension skills.

Canale (1984), after a consideration of the nature of receptive language proficiency, suggests that in order to assure an adequate representation of receptive language proficiency in a given test, minimal consideration be given to development stages, text demands, situations and purposes of language use, and their differing influences on the skills tested. He also brings to our attention the eight factors which influence language performance (Bachman 1987): (1) psychophysiological skills (sensory skills, and neuromuscular skills required); (2) representation of knowledge; (3) language use situations (reflexive or transitive; reciprocal or nonreciprocal); (4) context and message; (5) artificial restrictions (e.g. participants, mode, channel, and time/length); (6) monitoring factors; (7) affective factors; and (8) strategic factors. Finally, in his consideration of possible improvements in receptive language testing, he points out that a good test not only should reflect careful attention to validity (content, concurrent), reliability (internal, test-retest), and practicality (administration and scoring), but also should reflect test acceptability, and feedback potential. Hence, Canale offers two guiding principles for receptive language testing to counter the risk of overemphasizing psychometric and practical concerns in test development.

Principle 1: Attempt to elicit the best performance from test takers by presenting tasks that are fair, important in themselves, and interesting in themselves (bias for best).

Principle 2: Attempt to provide test tasks that can reveal to test takers and educators clear, rich, relevant, and generalizable information (bias for rewarding feedback).

Along with these principles, Canale mentioned four general test design features that are worthy of considerations in improving receptive language testing: (1) thematic organization; (2) four stages in test administration (warm-up, level check, probe stage, and wind-up); (3) adaptive testing procedures; and (4) criterion-referenced tests. Others have elaborated on these features. For example, Jones (1984) acknowledges the value of thematic organization by stating that "A thematic organization resembles more closely the types of listening and reading tasks we experience in real life," and that "in the construction of performance listening and reading tests, thematically organized passages are a necessary (p.366)". Nevertheless, he also points out the problems inherent in thematic test stimuli: they are less efficient; limit the variety of language materials that can be tested; and can result in test contamination. In a discussion of the 1986 ACTFL proficiency guidelines (American Council on the Teaching of Foreign Language), Douglas (1988) provides a brief overview of current work in listening test development and research. He suggests that current work in listening comprehension test development is moving in five directions: (1) of tests which are more efficient-adaptable to individual ability; (2) tests which are both integrative (in the technical sense of testing more than one skill or ability at once) and

integrated (in the sense of listening comprehension tasks feeding and being fed by other skills (e.g., the Ontario Test of ESL (Wesche 1987); (3) tests which challenge learners to employ varying strategies to deal with varying conditions of transmission noise, text mutilation or differing media; (4) tests which are contextualized; (5) tests which require students to apply understanding gained through listening to a variety of tasks involving other components of language proficiency (e.g., the task-based test design, Madden & Samuda 1985). According to Douglas there should be a continuing search for "a Grand Unification Theory in language testing that will relate constructs associated with language competence, domains of use, stages of learning, psychometric requirements and technological potential (p. 258)".

Basically, there are four stages in constructing a language test (Davies 1984):

Stage 1: The planning stage, planning the content and general layout of the test.

Stage 2: The prepilot stage, devising more items than will eventually be needed (preferably 3 times as many).

Stage 3: The pilot stage, trying out the test to check the administration procedure and provide materials for thorough item analysis and revision of the draft test.

Stage 4: The validation stage, trying out the test in its final form to obtain evidence as to its practical usefulness, reliability and validity.

According to Davies, stages 2, 3 and 4 are partly procedural and partly dependent on the success of stage 1. It is at this

stage that the professional language knowledge of the test constructor is called upon, and it is crucial for language sampling and hence for subsequent test validation. Therefore, for the sake of validating the new ESLPT Listening Comprehension Test, a detailed description of its content and general layout, as well as decisions made on the type of test items, will be provided in the later parts of the paper.

Part II: The New ESLPT Test Development

Stage 1: Planning Stage

Content and General Layout of the New Test

Since the new test is meant to be a new version of the present ESLPT Listening Comprehension Test, the format, final number of test items needed, and time spent in administering the test were all planned to be the same as for the present test. Thus, the new test was also to be a tape-recorded, multiple-choice format test, consisting of 3 parts, totaling 35 test items. The total administration time should be approximately 30 minutes.

The new test has taken a communicative language testing approach. The principle of authenticity is applied, and all the tasks undertaken in the test were planned to be as close to real-life, interactive communicative operations, and day-to-day discourse as possible.

The first procedure in developing the new ESLPT Listening Comprehension Test was to do a needs assessment to analyze the communicative needs of the participants. This was suggested for writers of communicative performance tests by Carroll (1980). As suggested by Richards (1985), were the situations, activities, and tasks in which our international students will be required to listen were analyzed. Apparently, academic listening is the main listening purpose of this specific group of students--ISU incoming international students--that we are dealing with. Based on our observation and literature surveys of related research, we developed a list of listening situations on the ISU campus and the activities involved. See Table 1.

TABLE 1 Listening Situations and Activities on Campus

SITUATION	ACTIVITY
In a Class	1. Listen to various types of lectures:
In a Recitation Section	narrative, description, process, instruction.
In a Lab Section	2. Listen to discussion.
At a Seminar	3. Listen to conversation.
At the Student Service Office:	1. Listen to introduction to services to be aware of the content of a service.
Student Counseling Office	2. Listen to schedule of events: when, where, what, how.
Health Center Office	3. Listen to procedures and fare of a service.
Administration's Office	4. Listen to directions to locate a place on campus.
Campus Information Desk	5. Listen to advice to properly manage a situation.
Library Service Office	
E.O. Building	
Women's Center	
Housing Office	
Telecommunication Office	
Media Resource Center	

Furthermore, a series of important listening skills were identified and selected as skills to be tested, in accordance with the ACTFL Proficiency Guidelines (1986) on listening competence at an advanced level (p.19), see Table 2a. And the results of Powers' survey of academic demands related to listening skills (1985), as well as Richards' taxonomy of academic listening skills (1985), see Table 2b.

TABLE 2a The ACTFL Proficiency Guidelines on Listening Skills at Advanced Level

Able to understand main ideas and most details of connected discourse on a variety of topics beyond the immediacy of the situation. Comprehension may be uneven due to a variety of linguistic and extralinguistic factors, among which topic familiarity is very prominent. These texts frequently involve description and narration in different time frames or aspects, such as present, nonpast, habitual, or imperfective. Texts may include interviews short lectures on familiar topics, and new items and reports primarily dealing with factual information. Listener is aware of cohesive devices but may not be able to use them to follow the sequence of thought in an oral text.

TABLE 2b Academic Listening Skills Selected for the New
Listening Comprehension Test.

(In accordance with Powers' survey of academic demands related to listening skills.)

1. Identifying major themes or ideas.
2. Identifying relationships among major ideas.
3. Identifying topic of a lecture.
4. Retaining information through note-taking.
5. Retrieving information from notes.
6. Inferring relationships between information.
7. Comprehending key vocabulary.
8. Deducing the meaning of words from context.
9. Following the spoken mode of lectures.
10. Identifying supporting ideas and examples.

(In accordance with Richards' taxonomy of academic listening skills.)

1. Ability to identify purpose and scope of lecture.
 2. Ability to identify topic of lecture and follow topic development.
 3. Ability to identify relationship among units within discourse.
 4. Ability to identify role of discourse markers in signaling structure of lecture.
 5. Ability to infer relationships (cause, effect, conclusion).
 6. Ability to recognize key lexical items related to subject/topic.
 7. Ability to deduce meanings of words from context.
 8. Ability to recognize markers of cohesion.
 9. Ability to recognize function of intonation to signal information structure.
 10. Ability to detect attitude of speaker toward subject matter.
 11. Ability to follow different modes of lecturing: spoken, audio, audio-visual.
 12. Familiarity with different styles of lecturing: formal, conversational, read, unplanned.
 13. Familiarity with different registers: written versus colloquial.
 14. Knowledge of classroom conventions.
 15. Ability to recognize instructional/learner tasks(e.g., warning, suggestions, recommendations, advice, instructions).
-

The situations and activities which were selected for the content of the new ESLPT Listening Comprehension test were those in which most students could easily picture their involvement. Those selected were:

1. Listening to conversations taking place in a library tour,

between a librarian and a student, regarding various services at the library (Part 1).

2. Listening to conversation taking place in a campus tour, between a tour guide and a student, regarding campus locations, and various campus services (Part 1).

3. Listening to classroom conversations, between a student and a teacher, regarding classroom policy, appointment, and assignment (Part 2).

4. Listening to lectures taking place in a class or a lab section regarding a variety of topics related to various academic fields.

The various listening skills mentioned by Powers and Richards were stressed in each part of the test. The skills focused on in Part 1 mainly were: "retaining information from notes", and "retrieving information from notes". In order to answer the questions in Part 1, the students need to be able to listen for very specific information on procedures, fares, and schedules, about various services on campus. Since in real life they might take notes on such information, they are permitted to do so in the test, thus demonstrating the ability to retain and retrieve information from notes. In Part 2, the student-teacher conversation part of the test, the focus was on testing students' knowledge of classroom conventions, and ability to recognize instructional/learner tasks (e.g., advice, suggestions, instructions.) In Part 3, the lecture part of the test, the focus was on testing the skills related to lecture comprehension, for example, "identifying major themes or idea, purpose and scope, as

well as topic of the lecture" and "ability to follow different modes of lecturing, spoken, audio, audio-visual" (both audio, and audio-visual modes were presented in the test). Though in each individual part of the test, some specific listening skills were the focus of testing, the other listening skills were not excluded. For instance, in the first and second parts of the test, the students still need to demonstrate the ability to recognize function of intonation to signal information structure. Besides, they often need to demonstrate the ability to decode the meaning of words from context, since very likely, they would not recognize every word in the texts.

Stage 2: Pre-pilot Stage

The Construction of scripts and Test Items

So far we have gone through Carroll's first phase - "design Phase" - of communicative language test construction, and the test content is specified. The next stage is the realization of design specifications phase, to write the test itself. As Carroll points out, this is perhaps the most difficult phase, since it requires a combination of scientific understanding and creative writing. The test writer is often expected to produce an attractive, workable, authentic set of tasks whose content and contextual features represent the situations in which English will be used. But often authenticity is likely to be abandoned in the search for reliability and economy. Therefore, to provide a test with some measure of authenticity, the selection of "source materials" is crucial.

For the situations and activities selected for the first part of the test, handouts and booklets related to various ISU

campus services were gathered and adapted. Final decisions were made according to the degree of popularity of each service, then scripts were drafted. The following campus services were chosen: library services, health services, and bus services. Also a genuine campus map was modified by a graduate student in Architecture, and a script was written, accompanying the map, for a tour around the campus. For Part 2 of the test, the scripts of the student-teacher conversations in a classroom setting were suggested by transcriptions of videotapes of real classroom interactions at ISU (Douglas and Myers 1987), along with the test writer's personal experience as a student. For Part 3 of the test, three lectures were included. The script of Lecture 1 was written by the test writer on a Sewage Disposal Unit, and a diagram was adapted from Patterns of Fact, an EAP testbook by Kennedy and Hunston (1982). For Lecture 2, on a Laser Jet Printer, both the script and diagram were adapted from a printer manual. For Lecture 3 on Technical Writing, both the script and bar chart were adapted from an EAP testbook on Technical Writing.

Efforts were made to ensure that the listening texts were distinct from the written texts, i.e. more colloquial (complex sentence structures were avoided, the rate of delivery was moderate, and intonation was natural). Nevertheless, since the script was read, some features of authentic communications were not included, such as background noise, false starts, pauses, and repetitions. In addition, the amount of information in each piece of discourse was reduced, so that it was reasonable to present within the time allowed. Jargon was avoided so that no discipline

would be favored. The selected academic listening skills to be tested were kept in mind while the scripts were written, or adapted. Since this is mainly a listening comprehension test, the written multiple-choice answers were either made short, or extra time was allowed for reading the longer answers to avoid involving too much reading skill in the test (See the script and testbook in Appendix A and B).

All together there were 45 items written for the scripts initially: 20 items for the first part, 8 for the second, and 17 for the third. The instructions given for answering each part of the test clearly informed the students of the specific listening purposes. For example, the instructions for Part 1 of the test informed the students that they would hear a series of conversations taking place during a library tour and a campus tour. They were told to listen carefully to specific information such as directions, procedures, fares and schedules. The students were to take notes on both Part 1 and Part 3, using the notetaking page provided or directly on the diagrams and bar chart, whenever there was a need.

The Taping

Careful procedures were taken to ensure that the test had good audio quality. The scripts and answers were taped at the Media Production Unit at ISU by professional technicians, and three experienced ESL teachers, a man and two women, recorded the script.

Stage 3: Pilot Stage

Subjects

After completion of the draft test, the 45-item, multiple-

choice listening comprehension test had its first try-out on 93 ESL students. They were students who were placed in English 100E classes by the present ESLPT listening comprehension test in August 1987. The students came from various language backgrounds and academic fields. The draft test was given to them in their final examination period in December 1987, after they had had a semester of independent study on academic listening skills. The reason for having this group of students as our subjects for the pilot testing was that they were similar to the kind of students on whom the test was to be used eventually.

Procedures and Problems in Administering the draft test

All the subjects were given two testbooks, one for answers, and the other containing a note-taking page, the campus map, two diagrams, and the bar chart. They were also given a computer scoring sheet, since in the concurrent testing which would be given in January 1988, with the present ESLPT Listening Comprehension Test, both tests would be scored by computer. The subjects needed to be able to handle two testbooks together, going back and forth from the testbook with notes and visual aids (map, diagram, bar chart) to the answer book. This was different from what is required in the present ESLPT Listening Comprehension test, where there is only an answer book to work on. This turned out to be not much of a problem, except that the test administrator stopped a few times after starting the tape to make sure that the subjects had found the right place, were looking at the right diagram, or were ready to go to the next part of the test.

Other than what is mentioned above, it was found that the instructions given for each part of the test seemed to serve the purpose; the subjects were able to follow the instructions given to do the task required. And the length of pauses allowing the subjects to read through the answers and put the chosen one down on the computer-scoring sheet seemed to be adequate as well. The whole administration time of the test was around 45 minutes.

Results and Comments

After the pilot testing, the answer sheets were scored, and raw scores were tabulated. Descriptive statistics, along with the reliability of each part of the test and the test as a whole were gathered. Table 3 contains the means, standard deviations, and reliabilities for the test.

TABLE 3 Mean, standard deviation, and reliability for New ESLPT Listening Comprehension Test.

Test	No. of Items	Mean	Standard Deviation	Reliability KR-21
Part 1	20	9.9	2.5	.24
Part 2	8	5.3	1.5	.30
Part 3	17	7.2	2.5	.38
Total	45	22.5	4.7	.50

With a reliability equal to or less than .50, the total test and the individual parts all proved to be unsatisfactory. But one may argue that with a group of subjects having similar listening proficiency level (placed together by the present ESLPT Listening Comprehension Test), low reliability is likely to occur. We were sure that this situation would improve in the concurrent testing, when there would be a bigger group with a wider range of

proficiency. Nevertheless with a mean of 22.5 (50%), the test was considered to be a hard test. Therefore, special effort needed to be made, such as changing the rate of delivery, or modifying distractors, to make the test easier.

TABLE 4 Correlation between the total score and parts. New ESLPT Listening Comprehension Test.

	Part 1	Part 2	Part 3	Total
Total	.77	.48	.77	1.00
All Coefficients $P < .001$				

Each part of the test was correlated with the total. These correlations are shown in Table 4. They were satisfactory, with only one exception: the correlation between the total score and Part 2. This lower correlation might be explained by a smaller number of test items. In Part 2 there were only 8 items, while in the other two parts, there were 20, and 17 items respectively. Also, the student-teacher conversations in Part 2 were less discourse oriented, having a shorter context in comparison to the other two parts, and this might be another reason for the lower correlations.

Item Analysis

An item analysis was conducted on each test item. For each item both "Facility Value" and "Discrimination Index" were calculated, the following formulas being used:

Formula 1: "Facility Value" equals the number of the top (higher score) one-third plus the bottom (lower score) one-third of the subjects who got the item correct, divided by the total number of the subjects in both groups.

$$FV = (1 + 3) / N_{(1 + 3)}$$

Formula 2: "Discrimination Value" equals to the number of the top one-third of the subjects with the right answer, minus the number of the bottom one-third of the subjects with the right answer, divided by the number of subjects in either group.

$$DV = (1 - 3) / N_{(1, 2)}$$

If an item has a "facility value" in the range between 30%-70%, it is considered satisfactory, i.e., neither too hard nor too easy; if it is below 30%, that means the item could be too difficult; if it is above 70%, the item could be too easy. As to the "discrimination index", if an item has a "discrimination index" over 40%, it is considered satisfactory, i.e., the item can discriminate the high proficiency level from the lower proficiency level; if it is below 40%, it indicates the item does not discriminate very well. See Table 5 for the "facility value" and "discrimination index" of each test item.

TABLE 5 "Facility Value" and "Discrimination Index" of pilot test items

Item No.	Facility Value	Discrimination Index
1	*75%	43%
2	37%	*13%
3	37%	*27%
4	*80%	*13%
5	70%	*33%
6	30%	*-13%
7	59%	*10%
8	*73%	*33%
9	*22%	*-17%
10	45%	*17%
11	52%	*30%
12	*18%	*10%
13	*23%	*27%
14	33%	*13%
15	60%	47%
16	37%	*20%
17	53%	40%
18	*97%	* 0%
19	62%	*37%
20	52%	50%
21	*77%	*33%
22	50%	*20%

23	45%	*37%
24	48%	*23%
25	*73%	40%
26	*80%	* 7%
27	*80%	*20%
28	69%	*33%
29	30%	*27%
30	52%	*23%
31	52%	57%
32	70%	*20%
33	48%	*30%
34	47%	40%
35	33%	40%
36	52%	43%
37	37%	*27%
38	45%	*-10%
39	48%	*30%
40	35%	*10%
41	30%	*33%
42	*20%	*-07%
43	47%	* 7%
44	42%	*23%
45	55%	*17%

"*": Items that are unsatisfactory and need to be revised or eliminated.

Results

1. The items with a minus or 0% "discrimination index" were eliminated, since these items would not serve the purpose of properly diagnosing the proficiency level of the ESL student. These items were: #6 (-13%), #9 (-17%), #18 (0%), #38 (-10%), #42 (-7%).

2. The items with a "facility value" that was either too high or too low were eliminated. These items were: #12 (18%), and #13 (23%).

3. A decision was made to omit the text of the campus map, and test items (#9-#14) related to it. Due to the fast rate of delivery, either low "facility value" or "discrimination index" was found in most of the items. In fact, only 1 item, #11, was adequate.

4. Some items with low "discrimination index" were revised.

Distractors were changed to be less attractive for the subjects at the higher proficiency level and more attractive for those at the lower level. These items were #2 (13%), #4 (13%), #7 (10%), #26 (7%), #27 (20%), #40 (10%). The changes in these items are as follows: (A star is placed in front of the right answer.)

Item #2: Which type of materials can ordinarily be obtained through interlibrary loan?

- *a. new books
- b. magazines
- c. encyclopedias
- d. videotapes

Distractor b - "Magazines" - was chosen by many (8) top group students, and was changed to "Periodicals," which was the exact word used in the listening text. The presumption for the higher proficiency level students, should be able to recognize the word, and therefore, make the necessary inference and select the right answer for the question. The lower proficiency level students should find the exact word "Periodicals" attractive, since their comprehension strategies might not be as good as those of the higher proficiency level students.

Item #4: What's the purpose of the Physical Science Reading room?

- a. Training faculty in physics.
- * b. Supporting graduate study in science.
- c. Encouraging students to study physics.
- d. Keeping a collection of materials.

Distractor a - "Training faculty in physics" - was not attractive to either the top or the bottom group students, and was changed to "Selecting science materials for students". The reason for the change was that "training faculty" did not appear anywhere in the listening text, but "choose" a synonym of "select" and "science

materials" were in the text. Again, the presumption was that the lower proficiency students would be attracted to it, since they might recognize the words appeared in the listening text, while the higher proficiency students would be more likely to comprehend the meaning in context and get the semantic information needed to provide the right answer.

Item #7: What is the first step in the search procedure?

- a. Study a search manual.
- b. Search through related topics.
- * c. Select a suitable search area.
- d. Expand topics with specific words.

Distractor a - "Study a search manual" - was changed to "Study a bibliography list". "Study a search manual" attracted a great many top students, probably they were torn between their world knowledge and listening proficiency, since often in the real world, people do study a manual first in learning to use a new technology product. To avoid the confusion, "Study a bibliography list" was used instead. Since "bibliography list" was mentioned in the listening text, it should be more attractive to students who do not comprehend the whole passage. But for those who do comprehend, the revised distractor should be less attractive.

Item #26: The teacher talking to his class is probably implying...

- a. His policy is relaxed.
- * b. They should attend every class.
- c. He has no preference.
- d. It is all right to be late.

Distractor a - "His policy is relaxed" - and d - "It is all right to be late" - were in a way similar. The former did attract some students, while the latter did not attract any. To avoid the similarity which created a useless distractor, "relaxed" was

changed to "fair", since "fair" seemed to be more commonly used when talking about "policy". And "It is all right to be late", "late" was changed to "absent", since "late" is not mentioned in the listening text, and in the conversation, the student is trying to find out about the teacher's attendance policy, students who only comprehend part of the passage, might relate "attendance" to "absence", therefore, the revised distractor should be more attractive.

Item #27: What would be an appropriate response for the student in this case?

- a. I'll see you at ten.
- b. It couldn't be better.
- * c. When would be a better time?
- d. Where should we meet?

Distractor d - "Where should we meet" - was not attractive to either the top or bottom students. If a student did not comprehend what the teacher was trying to suggest, that they meet at another time, the student would likely choose "I'll see you at ten" or "where should we meet" as his/her response. Since "see you" and "ten" were mentioned in the text, the students were more tempted to choose "I'll see you at ten" as the response, and pay no attention to the other incorrect response. Therefore, a decision was made to change it to "Really, I hope you can make it", to attract those students who could not tell the inappropriateness of the response in formality. The response would be more appropriate if the conversation was between two friends, but, in this case, it was between a teacher and his student. Furthermore, lower proficiency students may fail to see the importance of applying the cooperative principles when

carrying on a conversation. They may fail to realize that the teacher was implying that the students should ask him for an alternative, meeting him at another time. As for the higher proficiency level students, they should be able to apply the cooperative principles to carry on a conversation, providing an appropriate question.

Item #40: Which title do you think would be a proper title for the lecture?

- a. The body of a report
- b. How the readers read
- c. Different type of readers
- * d. Analyzing the reader's needs.

Distractor b - "How the readers read" - attracted many (13) top group students. Upon closer examination, there was indeed a similarity between it and the right answer "Analyzing the reader's needs". Therefore, a decision was made to change it to "Technical report writers" a supporting idea, rather than the main idea in the lecture. It was felt that the revised distractor would continue to attract the lower proficiency students, since they might not be able to differentiate main ideas from supporting ideas, while the higher proficiency students would be more drawn to the right answer for the question, since they no longer would have to choose between two similar right answers. (See Appendix C, the revised testbook. The distractors that underwent revision are underlined.)

Originally, the plan was to have the test administrator play the tape straight through, and that is what was planned to do in the concurrent testing. To avoid having to stop the tape in the middle of the test, a list of procedures for administering the test were constructed based on our experiences with the pilot

testing, included a clear explanation about how the visual aids worked with each part of the test.

In addition, the total administration time of 45 minutes was longer than what was aimed at for the final version of the test. However, after the revisions based on information gathered from the results of the pilot testing, only 35 test items remained in the test, and the time of the test was reduced to around 30 minutes. It was also expected that by giving concise explanations, unplanned stops would be eliminated, and some administration time could be saved as well.

Stage 4: The Validation Stage - Concurrent Testing

As was planned, the 35-item revised ESLPT Listening Comprehension Test was given concurrently with the present ESLPT Listening Comprehension Test to 151 incoming ISU international students on January 15, 1988. Based on the pilot testing experiences, a list of steps for administering the new ESLPT Listening Comprehension Test was compiled (see Appendix D), and was followed carefully by the test administrators, to minimize possible influence on the test results by administration practices. The 151 students were divided into two separate rooms: one containing 31 students with two administrators and the other containing 120 students with 5 administrators. The present listening test was given first, followed by the remaining multiple-choice parts of the ESLPT. The new listening test was given 5 minutes after the old ESLPT been given. In other words, the two listening tests were given approximately 50 minutes apart from each other. Students were not told that the new test was

experimental.

The purpose of giving the two tests concurrently was to validate the concurrent validity of the new test. Since the new test was designed to be a new version of the present ESLPT Listening Comprehension Test, the concurrent validity of the new test is of great importance. In addition than this, we were interested in finding out whether the test statistics and item analysis would show an improvement on our second try-out of the new test, and whether the preliminary revisions would be proved effective. In the following sections, a comparison between the test statistics of the two tests will be made. Correlations between the two tests, as well as findings regarding the effectiveness of the preliminary revisions, will be discussed.

Results and Comments

After the concurrent testing, both tests were immediately scored by computer. Descriptive statistics and reliabilities of the two tests were calculated (see Table 6).

TABLE 6 Descriptive statistics and reliabilities for the present and new ESLPT Listening Comprehension Test

Test		No. of Items	Mean	Standard Deviation	Reliability K-R21
Part 1	*P	8	4.34	1.78	0.60
	*N	12	6.93	2.11	0.47
Part 2	P	17	13.82	2.44	0.64
	N	8	5.82	1.59	0.47
Part 3	P	10	6.28	2.21	0.62
	N	15	7.12	3.10	0.68
Total Test	P	35	24.45	5.33	0.81
	N	35	19.87	5.53	0.78

* "P": The present ESLPT Listening Comprehension Test.

"N": The new ESLPT Listening Comprehension Test.

1. With a mean of 24.45 vs. 19.87, the present test was considered to be easier than the new test.

2. With a reliability of .81 vs. .78, the present test had a little higher reliability than the new test. But the reliability of the new test was satisfactory, and there was significant improvement in comparison to the reliability of the pilot test.

3. When making a comparison of the reliability between parts of the tests, for Part 1 and Part 2, the present test had a higher reliability, but for Part 3, the new test had a higher reliability.

Next, each part of the present test was paired with a part of the new test, based on the similarity of the type of listening task involved in each part of the two tests. Part 1 of the present test and Part 2 of the new test - comprehension of sentences - were paired together (pair 1), though Part 2 of the new test are short conversations, these two parts are considered to be more similar than the others. Part 2 of the present test was paired with Part 1 of the new test-comprehension of discourse conversations (pair 2), and Part 3 of the present test was paired with Part 3 of the new test-comprehension of lectures (pair 3). An example of each part of the present test is shown at the beginning of Part 1 of this paper, and an example of each part of the new test in the Appendix. Correlations were calculated between the two tests as a whole, the part and the total test score, and also between pairs with similar listening tasks. See Table 7 for the correlation coefficients.

TABLE 7 Correlation coefficients between Present ESLPT Listening Comprehension Test and the new ESLPT Listening Comprehension Test.

	PEPT	NEPT	PPART1	NPART2	PPART2	NPART1	PPART3	NPART3
PEPT	1.00							
NEPT	.73	1.00						
PPART1	.77	.60	1.00					
NPART2	.56	.67	.51	1.00				
PPART2	.86	.61	.54	.48	1.00			
NPART1	.56	.79	.48	.36	.51	1.00		
PPART3	.82	.59	.46	.40	.55	.42	1.00	
NPART3	.63	.89	.50	.44	.50	.54	.55	1.00

All Coefficients $P < .001$

1. A correlation of .73 indicated that the new test correlated fairly well with the present test. They both were probably testing a student's ability to comprehend listening texts.
2. With correlations of .77, .86, .82, each part of the present test correlated very well with the present test as a whole.
3. With correlations of .79, .67, .89, each part of the new test also correlated well with the new test as a whole with only one exception between Part 2 and the whole test. Although the correlation of .67 did show some improvement in comparison with that of the pilot test, it is still lower than desired.
4. Correlations of .51, .51, .55, the two parts within each pair were correlated, but not very highly. These results might be due to the content of the specific listening texts that were selected for each test. But no hard empirical evidence was gathered to support this assumption.

Item Analysis

An item analysis which was somewhat different from that done in the pilot test was conducted on each test item of both tests. Though the method and criteria used for the analysis here and the pilot test are varied, similar results are expected if both types of analysis were used. For each item "item difficulty" - the percentage of students answering the item correctly, as well as "item discrimination" - the item-score correlation were calculated. Items in the medium range of difficulty, about 30-70 percent students answering correctly, were considered satisfactory statistically. Items with extreme item difficulties, 0-29 or 71-100 percent, were considered to be either too easy or too difficult. As to the values of "item discrimination", negative values indicate an inverse relationship between total score and the score on the item. Such items therefore, should be eliminated or rewritten. Items with values falling in the range of .20-.39, were generally considered to be satisfactory, and items with values below .20 do not discriminate well among students, and those in the range of .40+ were considered as providing acceptable discrimination. See Tables 8 and 9 for the "item difficulty" and "item discrimination" of each test items on PESLPT and NESLPT respectively.

TABLE 8 "Item difficulty" and "Item discrimination" of the Test Items on PESLPT

Item No	Item Difficulty(%)	Item Discrimination
1	34	.29
2	39	.62
3	*90	.24
4	40	.48
5	*94	*.16
6	52	.49

7	62	.48
8	*23	.28
9	*79	.51
10	*71	.43
11	70	.50
12	57	.42
13	*85	.35
14	*87	.36
15	69	.39
16	*98	*.17
17	*96	.23
18	*85	.25
19	*95	*.16
20	*95	.36
21	*77	.45
22	*83	.23
23	*76	.34
24	*85	*.14
25	*74	.27
26	*81	.38
27	*74	.51
28	60	.55
29	*79	.38
30	68	.31
31	66	.46
32	52	.61
33	52	.21
34	34	.22
35	62	.29

 "*" : Items that are unsatisfactory.

TABLE 9 "Item difficulty" and "Item discrimination" of the Test Items on NESLPT

Item No.	Item Difficulty(%)	Item Discrimination
1	*83	.24
2	42	.30
3	31	.23
4	69	.41
5	63	.34
6	58	.27
7	61	.49
8	*79	.45
9	33	*.18
10	35	.32
11	69	.26
12	*73	.13
13	*79	.39
14	65	.46
15	54	.44
16	60	*-.12
17	*79	.36

18	*88	.27
19	*80	.49
20	*78	.26
21	30	.32
22	62	.41
23	51	.50
24	69	.29
25	49	.24
26	66	.40
27	32	.43
28	52	.53
29	47	.39
30	50	.38
31	36	.31
32	*21	.25
33	54	.47
34	47	.37
35	49	.44

"*": Items that are unsatisfactory.

Comments

1. Item difficulty. The present test had more easy items than the new test, i.e. there were 6 items, item #3, 5, 16, 17, 19, and 20, with item difficulty values above 90% on the present test, while there was none in this range on the new test. Also, on the present test, there were 20 items with item difficulty value above 70%, while on the new test only 8 were found. Item #8 on the present test, and item #32 on the new test were the only two with item difficulty below 30%. Therefore, in terms of item difficulty values, the new test was considered more satisfactory than the present one.

2. Item discrimination. Interestingly, on both tests, the same number of test items (13) with discrimination values above .40 were found on both tests, and the same number of test items (4) with discrimination values below .20 were found on both tests. But, item #16 on the new test was the only item with a negative discrimination value, $-.12$. Therefore, in terms of item

discrimination values, the present test was considered a little better than the new test.

3. The test items of the new test with distractors which had undergone revisions based on the pilot testing results all showed an improvement on their discrimination value. These items were Item #2, 4, 6, 18, 19, and 31, which were #2, 4, 7, 26, 27, and 40 on the pilot test.

Based on the results of the concurrent testing, the reliability, the correlation coefficients, and the item analysis, the revisions of the new test proved to be effective.

Subject and Language/Discipline Bias

Our tests were to be used for placing ESL post-admissions students entering the university for academic purposes. Although at this point, we already have discussed the reliability, correlations, and results of the item analysis of the two tests, there is still one question left to be answered:

Were the listening texts of the two tests neutral in content and cultural assumptions, so that they would not in any significant way favor any particular language group or academic field?

To answer the question of whether the tests were biased in favor of any particular language group empirical data were gathered and analyzed. The students were divided into 24 groups by their native language. The six largest groups were: Chinese (39), Indonesian (18), Urdu (17), Korean (13), Arabic (12), and Spanish (9). Eight SPSS-X analyses of variance were carried out comparing mean total and part scores on the two tests for the 24

groups and for the six largest groups. The results for all 24 groups indicated that there was no significant effect between language and the total score, nor between language and the score of each part of the tests. When the analysis was done for the six largest groups, there still were no significant differences among groups at the .05 level. The means and standard deviations of the scores of the six groups on both tests as a whole and the parts of the tests are presented in Table 10.

Table 10 Mean, Standard Deviation of the Six Largest Language Groups

Test/Group		Chin.	Indo.	Kore.	Urdu.	Span.	Arab.
PEPT	M	22.85	24.83	20.15	24.82	25.22	26.67
	SD	5.35	4.25	4.22	5.11	5.89	5.50
NEPT	M	19.08	20.17	16.92	20.76	18.11	20.83
	SD	5.64	5.04	4.50	5.33	6.93	5.11
PPART1	M	4.05	4.00	2.85	4.76	4.22	5.17
	SD	1.49	1.75	.99	1.75	2.82	2.00
NPART1	M	6.80	6.83	5.61	7.59	5.89	6.67
	SD	2.14	1.69	2.14	1.58	2.20	1.56
PPART2	M	13.31	14.28	12.00	13.76	14.33	14.42
	SD	2.54	1.64	2.64	1.92	2.60	2.97
NPART2	M	5.72	5.89	5.77	5.76	5.56	6.50
	SD	1.65	1.53	1.09	1.79	2.83	1.31
PPART3	M	5.49	6.56	5.31	6.29	6.67	7.08
	SD	2.33	1.98	1.60	2.44	1.73	1.51
PPART3	M	6.56	7.44	5.54	7.41	6.67	7.67
	SD	3.32	2.87	2.44	3.00	2.65	3.42

Subject and Discipline Bias

In order to answer the question of whether the tests were biased in favor of any particular discipline, the students were regrouped into four content areas according to their academic field. Urquhart and Alderson's method of grouping to assess the

effect of a student's discipline on a reading comprehension test (Hauptman 1985) was adopted here. For this analysis there were only 141 subjects, since 10 of those who took the tests did not indicate their discipline. These four groups were:

- Group 1: Development Administration, Finance, and Economics (DAFE), 16 students. Disciplines included: Business Administration science, Economics, Hotel Restaurant and Institution Management.
- Group 2: Engineers, 62 students. Disciplines included: Aerospace Engineering (Engr.), Agriculture Engr., Chemical Engr., Civil Engr., Electrical Engr. and Computer Engr., Industrial Engr., Material Science and Engr., Nuclear Engr.
- Group 3: Science and Mathematics, 35 students. Disciplines included: Agronomy, Animal Science, Computer Science, Earth Science, Engr. Science and Mechanics, Food Technology, Horticulture, Math, Microbiology, Plant Pathology, Seed and Weed Science, Professional Studies, Statistics, Veterinary Microbiology and Preventive Medication.
- Group 4: Liberal Arts, 28 students. Disciplines included: Agriculture Education, Architecture, Art and Design, Family Environment, History, Industrial Education and Technology, Journalism and Mass Communication, Psychology, Sociology and Anthropology.

Again, eight SPSS-X analyses of variance were carried out. The means and standard deviations of the scores of the four groups on both tests as a whole and the parts of the tests are presented in Table 11, followed by Table 12 which shows the F-ratio resulted from each analysis, since significant differences between the discipline groups on the total scores, and scores of the parts of both tests were indicated.

TABLE 11 Group Means and Standard Deviations for Each Test and the Parts

Test\Group		DAFE	Engineering	Science& Mathematics	Liberal Arts
PEPT	M	25.13	25.84	24.37	21.54
	SD	5.11	5.18	5.12	5.27
NEPT	M	19.50	21.66	19.60	16.71
	SD	5.01	5.08	5.93	5.05
PPART1	M	4.00	4.76	4.26	3.71
	SD	2.13	1.64	1.92	1.74
NPART1	M	6.94	7.34	7.17	5.57
	SD	1.61	2.07	2.01	1.95
PPART2	M	14.44	14.21	13.94	12.57
	SD	1.86	2.54	2.34	2.30
NPART2	M	6.13	6.02	5.91	5.14
	S	1.71	1.43	1.63	1.88
PPART3	M	6.69	6.87	6.17	5.25
	SD	1.89	2.18	2.15	2.24
NPART3	M	6.44	8.31	6.51	6.00
	SD	2.90	2.88	3.23	2.80

TABLE 12 F-ratios of Present and New Tests between Discipline groups

Test	F-ratios
PEPT	4.5322*
NEPT	5.7428*
PPART1	2.4920
NPART1	5.4022*
PPART2	3.5129*
NPART2	2.1889
PPART3	3.8600*
NPART3	5.4125*
* P < .05	

Table 12 shows that there were significant differences by major, except on Part 1 of the present test and Part 2 of the new test. And SCHEFFE analyses were done on all the discipline groups, but significant difference were only found between some

of the groups. Table 13 shows the result of the analyses, to show which groups were significantly different.

TABLE 13 Significant Differences between Liberal Arts and Other Discipline Groups on Mean Scores (SCHEFFE Analyses).

Test	DAFE	Engineering	Science & Mathematics
PEPT	N5	$p < .05$	N5
NEPT	N5	$p < .05$	N5
PPART2	N5	$p < .05$	N5
NPART1	N5	$p < .05$	$p < .05$
PPART3	N5	$p < .05$	N5
NPART3	N5	$p < .05$	N5

Comments

1. For both tests as a whole, significant differences at the .05 level were found only between the Engineering and the Liberal Arts groups.

2. For Part 2 of the present test, and Part 1 of the new test, significant differences at the .05 level were again found between the Engineering group and the Liberal Arts group. As mentioned earlier in the correlation analysis, these two parts were paired because they involved similar listening tasks. Additionally for Part 1 of the new test, significant differences were also found between the Science and Mathematics group and the Liberal Arts group.

3. For the lecture parts, Part 3 of both tests, once again significant differences were found between the Engineering group and the Liberal Arts group. But for Part 3 of the new test, significant difference was also found between the Engineering group and the Science and Mathematics group.

These findings can be interpreted in two different ways. One is that the content of both tests, were biased toward certain disciplines. The Engineering students tended to score significantly better than the Liberal Arts students on the tests. However, one may interpret the results differently. The finding may reflect a true picture of the English proficiency level of the students. In general, students entering into engineering-related fields are more competitive, and a higher standard is required for admissions. Therefore, they have the potential to do better on the tests. More empirical evidence needs to be gathered before any assertion can be made.

Part III: Conclusions and Suggestions for Further Revisions

Conclusions

In this paper, a new ESL post-admissions listening comprehension test has been discussed, including its theoretical bases as well as its developmental and validation stages. A comparison was made between the current ESL listening comprehension test for placement purpose at ISU and the new test. The results are summarized in the earlier parts of the paper. Following are some of the conclusions drawn from the results:

1. The new test was developed in accordance with the current theories of language teaching and testing: the communicative approach, use of authentic language, and the teaching of language for specific purposes. Therefore, in terms of construct validity, the new test can be justified on the basis of theory.

2. The new test is acceptable in terms of its reliability and concurrent validity. This conclusion is based on the fact that the total score of the new test correlates well with that of the current ESL Listening Comprehension Test.

3. The new test is not biased toward any language group, and despite the fact that significant differences were found between the discipline groups, further empirical data would need to be gathered (i.e. TOEFL scores of the students, to see if there are significant differences among discipline groups) before any assertions can be made that the differences were not rendered by any other factors than the test itself.

4. The new test is not a perfect test, there is indeed room left for further revisions.

Suggestions for Revision

1. The content of texts. Most of the texts were adapted from written materials. Though efforts were made to make them more like spoken language, they are still not distinct from the written language. It is recommended that the texts be modified, adding more features of the spoken language.

2. The intonation and rate of delivery. Though the intonation and rate of delivery was natural and adequate, the fact that the script was read, rather than delivered spontaneously, detracted from its authenticity. It is recommended that the scripts not be read, but rather serve as guide to the speakers.

3. The extra testbook. The extra testbook which contains the notetaking pages and the diagrams has made both the tasks of giving and taking the test somewhat more complicated. However, two testbooks are required in this test. It is suggested that more time be given whenever there is a need for flipping pages, and locating diagrams.

4. Bias toward certain disciplines. To avoid bias toward certain disciplines, it is recommended that, especially for the lecture part of the test, more general texts be selected for the purpose.

5. The improvement of reliability. Though the reliability for the whole test is satisfactory (.78), the reliability of Part 1 and Part 2 of the test are somewhat lower. It is recommended that those more difficult items can be revised either by modifying the distractor or changing the question.

6. The attention to "purpose" in EAP texts. Though the test

writer intended to select the texts in accordance with "purpose" in EAP texts, but this was not done consistently throughout the test. It is recommended that texts more related to academic purposes can be collected for further revisions.

7. The inferential meaning in questions. Though there are some inferential questions in the test, more questions tend to be testing the student's ability to recall information mentioned in the texts. In real academic listening, the student are required to do a lot of inferencing from the information they listen to. It is recommended that some test items can be revised to deal with inferential meaning.

BIBLIOGRAPHY

- Aiken, Edwin. et al. 1975. Memory for a lecture: effect of notes, lecture rate, and information density. JOURNAL OF EDUCATION PSYCHOLOGY 67.3: 439-444.
- Alderson, J. C., and A. H. Urquhart. 1985. This test is unfair: I'm not an economist. in P.C. Hauptman, et al, eds. SECOND LANGUAGE PERFORMANCE TESTING. Ottawa: University of Ottawa Press.
- Alderson, J. C., and A. H. Urquhart. 1985. The effect of students' academic discipline on their performance on ESP reading tests. LANGUAGE TESTING 2.2: 192-204.
- Bachman, L. F. 1988. Problem in examining the validity of the ACTFL oral proficiency interview. STUDIES IN SECOND LANGUAGE ACQUISITION 10.2: 149-164.
- Benson, C. P. et al. 1980. Listening competence: a prerequisite to communication. in J. W. Oller and K. Perkins, eds. RESEARCH IN LANGUAGE TESTING. Rowley MA.: Newbury House.
- Brown, H. D. 1987. PRINCIPLES OF LANGUAGE LEARNING AND TEACHING. NJ.: Prentice-Hall.
- Brown, G. 1978. Understanding spoken language. TESOL Quarterly 12.3: 271-283.
- Burbidge, B. 1984. Looking at volcanoes and listening comprehension. NYS TESOL Conference.
- Canale, M. 1984. Considerations in the testing of reading and listening proficiency. FOREIGN LANGUAGE ANNALS 17.4: 349-357.
- Canale, M. et al. 1984. The testing of reading and listening proficiency: a synthesis. FOREIGN LANGUAGE ANNALS 17.4: 389-392.
- Canale, M. and M. Swain. 1980. Theoretical bases of communicative approaches to second language teaching and testing. APPLIED LINGUISTICS 1.1: 1-47.
- Candlin, C. N. 1986. Explaining communicative competence: limits of testability? in C. W. Stansfield, eds. Toward Communicative Competence Testing: Proceedings of the Second TOEFL Invitational Conference. Princeton: Education Testing Service.
- Carroll, B. J. 1980. TESTING COMMUNICATIVE PERFORMANCE. Oxford: Pergamon Press.
- Chaudron, C. and J. Richards. 1986. The effects of discourse

markers on the comprehension of lectures. APPLIED LINGUISTICS 7.2: 113-117.

Davies, A. 1984. Validating three tests of English language proficiency. LANGUAGE TESTING 1.1: 50-69.

Davies, N.F. 1980. Putting receptive skills first: an experiment in sequencing. CANADIAN MODERN LANGUAGE REVIEW 36: 461-467.

Douglas, D. 1988. Testing listening comprehension in the context of the ACTFL proficiency guidelines. STUDIES IN SECOND LANGUAGE ACQUISITION 10.2: 245-260.

Dunkel, P. A. 1986. Developing listening fluency in L2: theoretical principles and pedagogical considerations. MODERN LANGUAGE JOURNAL 70.2: 99-106.

Faerch, C., & Cziko, G. 1985. Determining the reliability, validity and scalability of the graduated dictation test. LANGUAGE LEARNING 35: 555-566.

Finocchiaro, M. and S. Sako. 1983. FOREIGN LANGUAGE TESTING: A PRACTICAL APPROACH. New York: Regents Publishing Co.

Ganske, L. 1981. Note-taking: a significant and integral part of learning environments. ECTJ 29.3: 155-175.

Halliday, M. A. K. 1979. LANGUAGE AS SOCIAL SEMIOTIC. London: Arnold (publishers) Ltd.

Hutchinson, T. and A. Waters. 1987. ENGLISH FOR SPECIFIC PURPOSES. Cambridge: Cambridge University Press.

Jones, R. L. 1984. Testing receptive skills: some basic considerations. FOREIGN LANGUAGE ANNALS, 17.4: 365-367.

Kennedy, J. and S. Hunston. 1982. PATTERNS OF FACT. London: Edward Arnold (publishers) Ltd.

Liskin-Gasparro J. E. 1984. Practical considerations in receptive skills testing. FOREIGN LANGUAGE ANNALS 17.4: 361-364.

Mackay, R. and A. J. Mountford eds. 1978. ENGLISH FOR SPECIFIC PURPOSES. London: Longman Group Limited.

Madden, C. G., & Samuda, V. 1985. Task-based test design. LANGUAGE TESTING UPDATE 1: 11-15.

Oller, J. W. 1974. Expectancy for successive elements: key ingredients to language use. FOREIGN LANGUAGE ANNALS 7: 443-52.

Oller, J. W. 1986. Communication theory and testing: what and

how. in C. W. Stansfield eds. Toward Communicative Competence Testing: Proceedings of the Second TOEFL Invitational Conference.

Olshtain E. Crosscultural pragmatic and the testing of communicative competence. LANGUAGE TESTING.

Palmer, A. S. and P. J. M. Groot, G. A. Trosper eds. 1981. The Construct Validation of Tests of Communicative Competence. Washington, DC.: TESOL Conference.

Peck, M. L. and M. J. Hannafin. 1983. The effects of notetaking pretraining and the reminding of notes on the retention of aural instruction. JOURNAL OF EDUCATION RESEARCH 77.2: 100-107.

Peters, D. L. 1972. Effects of notetaking and rate of presentation on short-term objective test performance. JOURNAL OF EDUCATIONAL PSYCHOLOGY 63.3: 276-280.

Porter, D. and J. Roberts. 1981. Authentic listening activities. ENGLISH LANGUAGE TEACHING JOURNAL 36.1.

Postovsky, V. A. 1974. Effects of delay in oral practice at the beginning second language learning. MODERN LANGUAGE JOURNAL 58:229-39.

Powers, D. E. 1985. A survey of academic demands related to listening skills. Princeton: Educational Testing Service.

Richard, J. C. 1985. Listening comprehension: approach, design, and procedure. in THE CONTEXT OF LANGUAGE TEACHING, pp. 189-207. Cambridge: Cambridge University Press.

Seliger, H. W. 1985. Testing authentic language: the problem of meaning. LANGUAGE TESTING 2.1: 1-15.

Shohamy, E. and T. Reves. 1985. Authentic language tests: where from and where to? LANGUAGE TESTING 2.1: 48-59.

Skehan, P. 1984. Issues in the testing of English for specific purposes. LANGUAGE TESTING 1.2: 202-220.

Snow, B. G. and K. Perkins. 1979. The teaching of listening comprehension and communication activities. TESOL Quarterly 13.1: 51-63.

Spolsky, B. 1985. The limits of authenticity in language testing. LANGUAGE TESTING 2.1: 31-40.

Vesta, F. J. and G. s. Gray. 1973. Listening and notetaking. JOURNAL OF EDUCATIONAL PSYCHOLOGY 64.3: 278-287.

Wesche, M. B. 1987. Second language performance testing: the

Ontario test of ESL an example. LANGUAGE TESTING 4.1:
28-47.

Widdowson, H. G. 1983. Learning purpose and language use. Oxford:
Oxford University Press.

Wilkes, H. 1981. The aural component: what the teacher needs to
know. CANADIAN MODERN LANGUAGE REVIEW 38: 68-80.

Wing, B. eds. 1986. LISTENING, READING, AND WRITING ANALYSIS AND
APPLICATION. Northeast Conference on the Teaching of foreign
Languages.

Wipf, J. A. 1984. Strategies for teaching second language
listening comprehension. FOREIGN LANGUAGE ANNALS 17.4:
345-348.

(APPENDIX A - SCRIPT OF THE NEW ESLPT)

A TEST OF ACADEMIC LISTENING SKILLS

PART 1: Library Tour and Campus Tour

INSTRUCTIONS: You will hear a series of conversations taking place during a library tour and a campus tour. Listen carefully to specific information such as: directions, procedures, fares and schedules, you may take notes on the pages provided. You will be asked questions about the information later on.

SITUATION I (Library Tour-conversation between a librarian and students)

L: Here's our Interlibrary Loan Office. Let me tell you about the services that are available here. Interlibrary loan allows you access to library materials not held by our university library. All you need to do is fill out a request form, wait about 2-3 weeks, and you will be notified by mail when the material is received. You need to come as soon as possible after notification to pick it up, and you will be asked to sign a circulation slip to obtain the material.

S: Does it mean we can obtain any type of material through interlibrary loan?

L: Well, not exactly. Some types of material can't ordinarily be obtained on loan. For example, periodicals, reference books, very old or rare books, and audio-visual items.

S: How long can we keep the material?

L: Well, the loan period is usually 2 weeks from the date the material is received, unless you arrange for a renewal.

Borrowed materials must be returned to this office on or before

the due date. If you need to keep the material longer, you should request a renewal at least two days before the due date. One extension is usually possible, although some libraries will not renew loans. If the lending library refuses the extension, you will be notified, and the material must be returned right away.

Now, answer the following questions by marking the letter of the best choice on your answer sheet beginning with number 101. You must listen to the questions carefully because they are not written in your test book.

101. Which of the following is included in the procedures for interlibrary loan?

102. Which type of materials can ordinarily be obtained through interlibrary loan?

103. What is true about renewing borrowed materials?

Now, listen to the next part of the library tour.

L: Here's the physical Science Reading room. It contains a specialized collection of science materials chosen to provide support for graduate study.

S: Do they have the same schedule as the general library hours?

L: No, it's a little different. Monday to Thursday they are open from 8 in the morning to 10 in the evening; Friday they are open at the same time but close at 5 in the afternoon. As for the weekend, on Saturday they are open from 10 to 5; Sunday they don't open until 2 o'clock in the afternoon, but they close at 10 in the evening.

Now, answer the following questions beginning with number 104.

104. What's the purpose of the Physical Science Reading room?

105. When can the reading room be used?

Now, listen to the next part of the tour.

L: Now let me show you the Searchline Desks. They provide a library research service linking the computer, long distance communication equipment and many bibliography lists. You can use this service to find references quickly for a research topic.

S: Do we get somebody to do the research for us?

L: No, this is a do-it-yourself service, although using a computer bibliography does mean you'll have to learn the proper procedure, which you'll have to practice a bit.

S: How do you do it?

L: OK., first you'll have to choose the area best suited to your topic, identifying the most important concepts in your research topic, expanding and combining them with the words-"OR", "AND", "NOT". We have a handout to explain the procedure in detail, and you can pick one up from the shelves on your way out.

S: Thanks. By the way, is it a free service, or do we need to pay for it?

L: Well, there is a fee, it's 50 cents per minute of connected time, and \$2.50 for every 50 citations. It has gone up a little bit, it used to be 40 cents per minute.

Now, answer the following questions beginning with number 106.

106. Which sentence about Searchline is correct?

107. what is the first step in the search procedure?

108. How much does it cost for 3-minutes of connected time and 100 citations?

SITUATION II (Campus Tour---conversation between a tour guide and students). As you listen, look at the campus map in your test book. You may write notes on the map. You will be asked questions about it later.

G: Before going on our campus tour, let's look at the campus map together for a while; I'll show you where the key offices are, these are the places that we are going to visit later. First, the memorial library where we are now is located at pretty much in the center of the campus. Right across North Road, facing the library is Grand North Hall, both the Graduate School Office and the Admissions Office are in this building. The Treasurer's Office, where you need to go to pay fees and tuition, is in the Public Administration Building right behind Grand North Hall. The International Student Office, where you need to visit if you have questions about immigration law and visas, is located at a building further north. For us to get there, we need to follow North Road, past the University Health Center on our left. After crossing Broad Way we will see a Visitor's Parking Lot on the right; the office is in the building right next to the parking lot. If you need any advice on your study skills, the Student Counseling Office is in the same building.

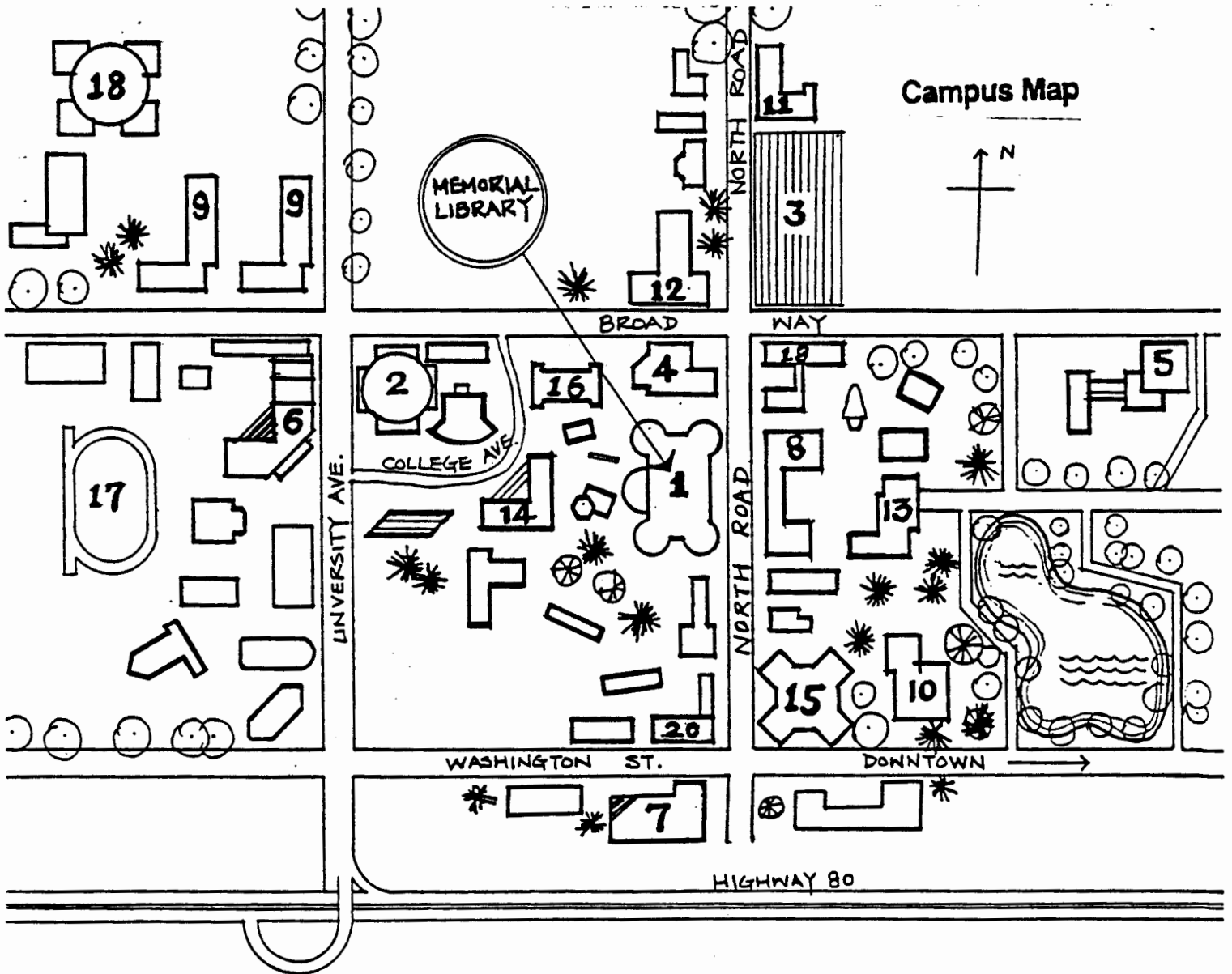
S: Will you also show us where the university bookstore is?

G: Sure, it's located in the Student Center Building. We'll need to go south, following North Road, turn right when we get to Washington street, and it's the first building on the left. By the way, there's a nice cafeteria too, just so you know where to go for food. Our final stop is at the Field House where you can checkout sport equipment and get some exercise. It's on the west

side of the campus, at the southeastern corner of Broad Way and University Avenue. For those of you living in the dorms on the northwestern corner of the two streets, it will only take you a couple of minutes to walk over, so you really should take advantage of the good facilities there. On our way to the Field House following Broad Way, we will make a left turn at College avenue, and stop by the second building on the left to visit the department of English where most of you will be taking some classes. Now we are ready to go for our trip, and as we go I'll show you some more buildings where different departments are.

Now, answer the following questions beginning with number 109 in your test book.

Campus Map



109. Which key office is in the Public Administration Building?

110. What number is the building where a student can go to get some lunch?

111. Which building is labeled #9?

112. What number is the building where a student can go for some academic counseling?

113. What number is the building where a student goes for English classes?

114. Which building is labeled #4?

Now, listen as the tour guide tells the students about the Health Center. You may take notes.

G: This is the Student Health Center. The university provides medical service for any enrolled student at a reduced rate. Services here include: the outpatient clinic, lab services, and physical therapy.

S: How do we pay for the services we receive at the Center?

G: Well, mainly there is a voluntary health fee. It's \$22 for a single student and \$50 for a student and spouse each semester, but there is an additional \$50 fee for prenatal care.

S: Does that mean I don't need to pay for any medical treatment if I paid the voluntary health fee?

G: Not quite, you'll also have to pay for services normally covered by medical insurance policies such as lab and professional fees resulting from accidents, lab tests and exams processed outside the Health Center, as well as prescriptions and pharmacy supplies.

Now, answer the following questions beginning with number 115 in your test book.

115. Which one of the following is available at the Health Center?

116. What is the total cost for a student family which receives prenatal care?

117. Which statement about the health fee is correct?

Now, listen as the guide tells the students about bus service. You may take notes.

G: Here's a bus stop sign, you'll find a lot of bus stops on campus. To ride on a bus you need to wait by a bus stop sign. You'll need to have the exact fare, or a ticket, or pass ready before boarding the bus. As a university student, you can show the driver your current fee card too, and ride for a reduced fare. The bus has different routes that overlap, you can go almost anywhere in town.

S: What should I do if I need to use two buses to complete my trip?

G: Yes, you will need a transfer, and you should ask for your transfer when you get on the bus, because the driver won't issue one at any other time. Transfers are only valid for the completion of a one-way trip, and you have to get on the first possible connecting bus.

S: Does the company have door-to-door service?

G: Yes they do. Dial-A-Ride is the name of their door-to-door service. This service is available to the elderly, handicapped, and the general public as well. Those buses of this special service are equipped with wheelchair lifts to serve the handicapped riders, and an attendant can ride free when accompanying a handicapped person.

Now, answer the following questions beginning with number 118 in your test book.

118. What does a university student need to have to ride for a reduced fare?

119. Which statement is true about transfers?

120. Who can ride Dial-A-Ride for free?

Now, get ready to listen to the instructions for part 2 of the test.

PART 2: Student-Teacher conversations.

INSTRUCTIONS: You will hear a series of conversations taking place in a classroom between a teacher and a student. Listen carefully. You will be asked about the meaning of the conversation, or about a proper response. Choose the best answer by marking the letter on your answer sheet. First, listen to an example question:

You will hear...

S: Do we need to memorize this equation?

T: I hope not. I'll talk about it a little more, I want to make sure you understand it, because once you understand you won't need to memorize it.

What is the teacher's intention by giving such a response?

you will read...

- a. giving a command
- b. giving advice
- c. making a request
- d. making a complaint

The best answer to the example is B giving advice.

Now, listen to the first conversation and answer Question 121 in your test book. The questions are not in the test book so listen carefully.

S: I still don't quite understand your points here.

T: Why don't you come to see me at my office sometime and we will talk it over.

121. What's the teacher's intention by giving such a response?

Conversation 2

S: You aren't going to cover chapter 9 in our final, are you?

T: We have talked about that chapter in class, haven't we?

122. What will be a proper response for the student?

Conversation 3

S: I haven't got my project ready yet.

T: I thought all of you were turning in your assignment today.

You really should have arranged to start your project earlier.

123. The teacher speaking to a student is probably...

Conversation 4

T: Susan could you rewrite your paper and turn it in by tomorrow?

124. The teacher speaking to a student is probably...

Conversation 5

T: If you have any questions feel free to stop me anytime during the lecture, or you might want to talk with me after the class.

125. The teacher talking to his class is probably suggesting...

Conversation 6

S: What's your attendance policy?

T: Well, let's just say I would like to see you in class everytime.

126. The teacher talking to his class is probably implying...

Conversation 7

S: I am wondering when would be a good time to see you tomorrow, is ten in the morning OK for you?

T: I really have a lot of things going on tomorrow morning.

127. What would be an appropriate response for the student in this case?

Conversation 8

S: Do you give pop quizzes?

T: Well, it's really not a bad idea.

128. The teacher talking to a class is probably implying that...

Now, get ready to listen to the instructions for Part 3 of
the test.

PART 3: Lectures

INSTRUCTIONS: You will hear some short lectures taking place in a classroom or a lab. Listen carefully to main ideas, supporting ideas, and specific information about processes and functions. You may take notes on the pages provided in your test book. After you have listened to the lectures, you will be asked some questions about them.

Now, listen to lecture I on the Sewage Disposal Unit. Look at the diagram in your test book. You may write notes on the diagram.

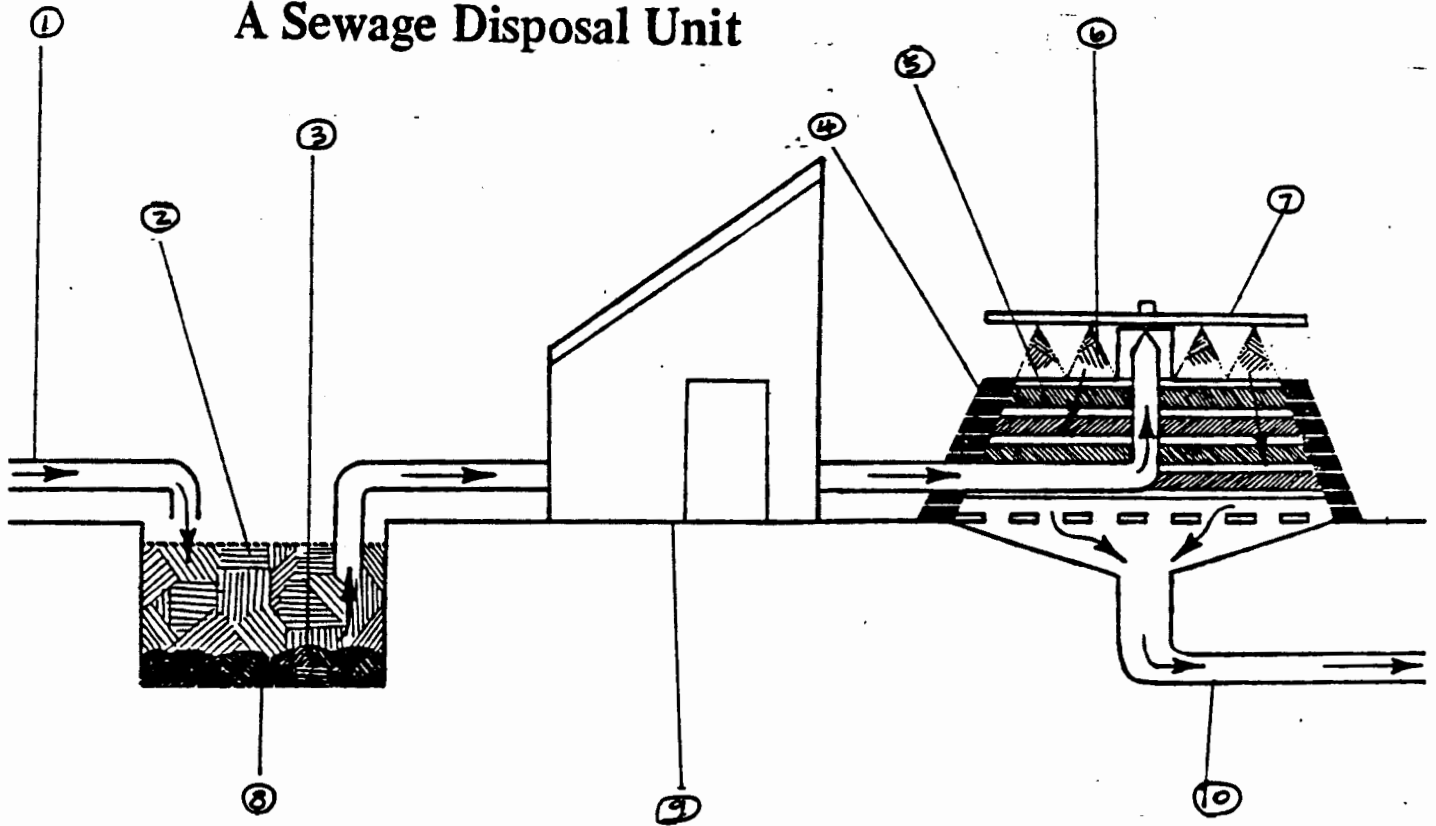
A Sewage Disposal Unit is essential for water purification. Look at the diagram. There are three main parts of the unit: the settling tank, the pump house, and the filter beds. They are connected to each other by a pipe with the pump house in the middle.

When polluted water flows into the settling tank from the main sewer, the undecayed solids sink to the bottom of the tank to be removed and burned later on. The water is pumped up through the pump house, and goes to the filter bed.

The best way for the polluted water to be purified thoroughly is for a smaller quantity of water to go through the filters at one time. So, when the water gets to the beds, it goes to the rotating arms, which sprinkle jets of polluted water down onto the filters of porous stones and micro-organisms. Once the polluted water has finished its trip through the sewage disposal unit, it's purified and ready to empty into a river.

Now, answer the following questions about the lecture,
beginning with number 129.

A Sewage Disposal Unit



129. Which arrow is pointing to the main sewer?

130. The rotating arm is labeled #....?

131. What is the first stage of water purification?

132. The part labeled #5 is...?

133. What is the reason for using rotating arms?

Now, listen to lecture II on the Laser Jet Printer. Look at the diagram in your test book. You may write notes on the diagram.

Laser Jet printers for computers are gaining in popularity now-a-days. In order for you to start using a Laser Jet printer, there are several important features that you need to know about. Look at the diagram. First of all, the Operation Control Panel is located on the top left portion of the front side, and is used to control printer functions such as choosing on-line (connected to the computer) and off-line (disconnected from the computer), and selecting a manual feed mode that allows you to feed in odd-size paper and envelopes through the Manual Feed Tray located on the rear side of the printer. In addition, the Status Display on the panel displays printer status information and error codes by showing a 2-digit number. For example, zero-zero indicates the printer is ready; "11" means you need to add some paper. The Paper Input Cassette is located at the bottom of the front side right underneath the Print Tray where you collect your printed paper. Occasionally, you might have number "13" which tells you a paper jam has occurred, and you will need to open the upper main body of the printer to check for jammed paper. The Upper Unit Release Level is located on the right side toward the front.

One of the useful features of a Laser Jet printer is its ability to accept optional font cartridges containing a variety of print styles. The Font Cartridge Slot is located next to the Operator Control Panel. To insert a font cartridge you need to, first, set the printer off-line, then slide the cartridge into the slot. After you get done with it, simply set the printer off-line again, grasp the edge of the cartridge and slowly pull it out of the slot, then press the on-line key to reconnect to

the computer.

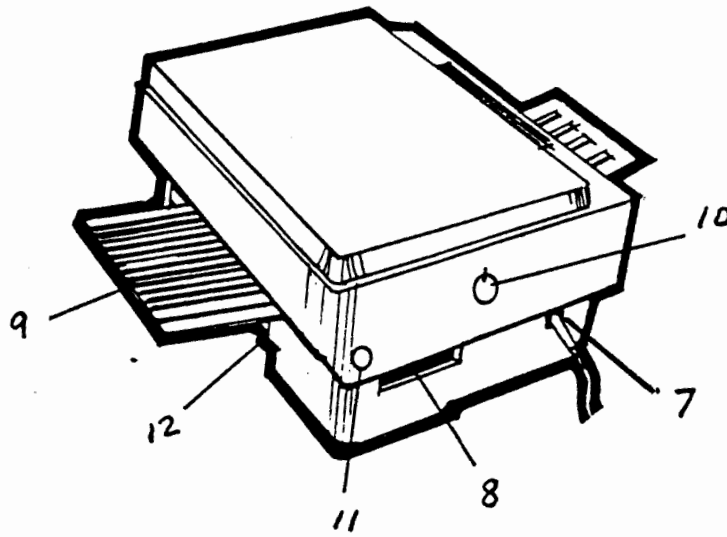
If you want to print data from a computer system, you will need to plug the interface cable into the Interface Connector at the lower part of the left side toward the rear. Also connect the AC power cord to the Power Cord Receptacle on the same side toward the front. When the printer is ready, press the on-line key, and it's ready to go to work.

If you wish to adjust the print darkness upon seeing the first printout, the Print Density Dial is located at the center of the upper part of the left side. You can use it to adjust the darkness of your printout by turning it clockwise to decrease the print density or counterclockwise to increase the print density.

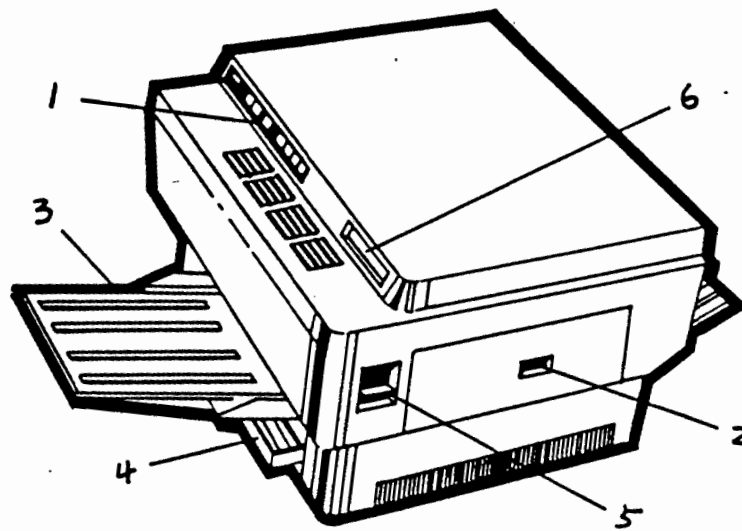
In case of any problem, you can always look in the "correcting error codes" portion of the maintenance and trouble shooting section of the manual, since it contains a detailed description of what to do for each error condition.

Now, answer the following questions about the lecture beginning with number 134.

Lecture II: The Laser Jet Printer



Printer Rear and Left Side



Printer Front and Right Side

134. Which arrow is pointing to the AC power cord receptacle?
135. Where do you feed in nonstandard paper and envelopes?
136. What is the first thing to do if you want to insert a font cartridge?
137. What do you need to do to print data from a computer?
138. What should you do if your printout is too light?
139. According to the lecture what is the first thing you should do if your printer isn't operating properly?
-

Now, listen to lecture 3 about Technical Writing. Look at the graph in your test book. You may write notes on the graph.

Not all technical report readers want or need the same information, so no piece of technical writing is universally useful. The most frequent mistake made by a novice technical writer, also the most serious one, is treating all audiences as though they were the same. The audiences of technical reports may include researchers, technicians, administrators, and laypeople, or it may be a variety of people with many different backgrounds and areas of expertise. The researchers and technicians are interested almost exclusively in their own disciplines, while what's more important for administrators is to get the specific information that they need within a short period of time, such as: how to further the organization's goals, how to increase the profits, and so on. They don't expect to learn anything about management from the technical writer, and they don't want to learn any more than they have to about the writer's technical field. But don't get it wrong, the trick to writing technical reports for administrators isn't condensation, it's selection.

Look at the bar chart. According to the result of a survey done in 1986 at Focus Electric Corporation on "how administrators read reports", technical writers who zero in on the body of a report to be read by administrators may find themselves wasting their efforts. The bodies of most technical reports simply take more time to read than many administrators have to spare. The results show that the body of the technical report was read only 28% of the time by administrators. By contrast, the introduction was read 70% of the time and the conclusion 64%. The most read section of technical reports was the summary section. The

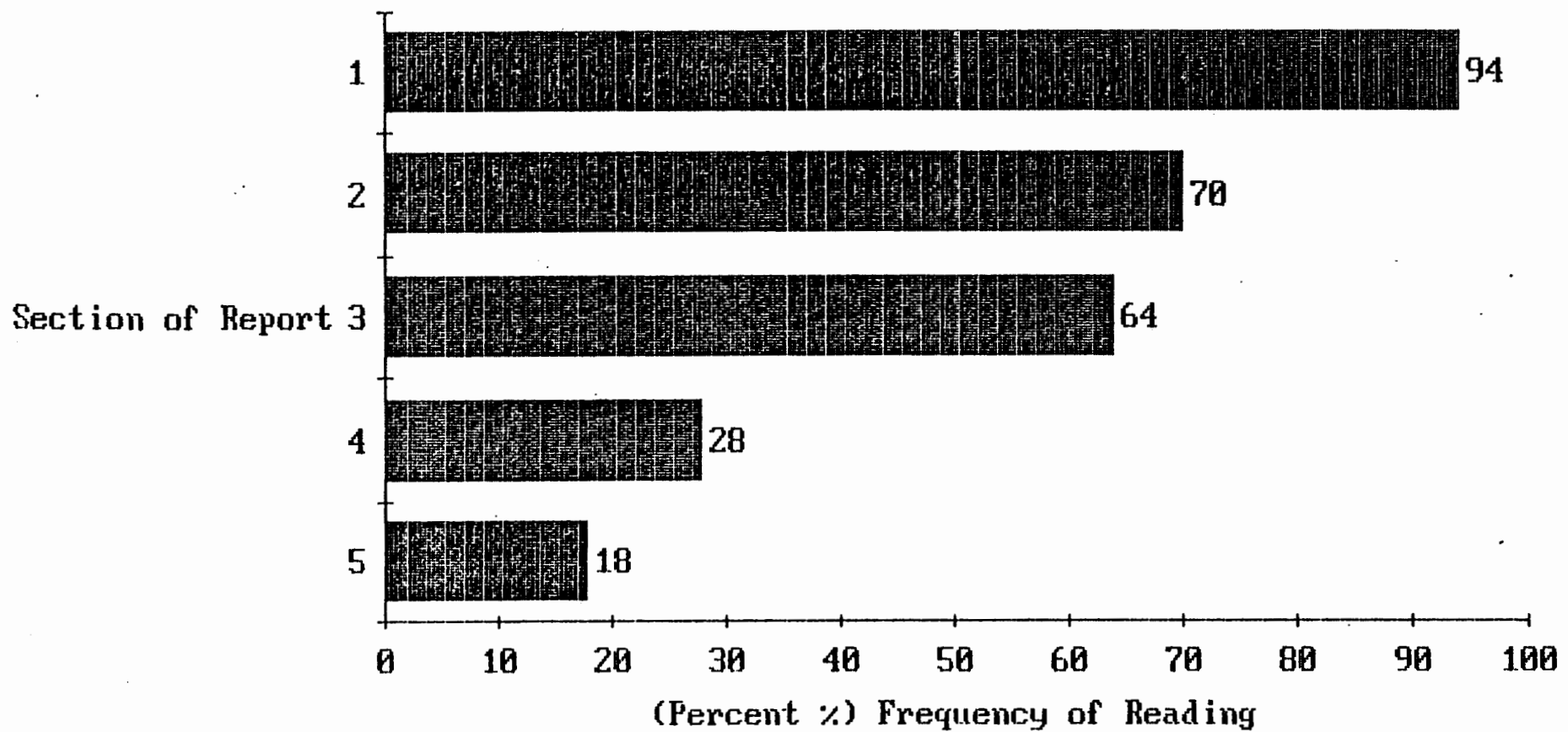
appendices were read only 18% of the time by administrators.

As a writer of technical reports, keeping your readers' concerns in mind while writing a report, will help you to include the information needed to make your reader's job more efficient and yours more effective.

Now, answer the following questions about the lecture, beginning with number 140.

Lecture III: Technical Writing

Figure 1.1 How Administrators Read Reports



140. Which title do you think would be a proper title for this lecture?

141. What type of mistake is a new technical writer most likely to make?

142. What kind of technical report does an administrator prefer to read about?

143. Which section of a technical report is labeled #1 on the bar chart?

144. Which section of a report is labeled #5 on the bar chart?

145. Which section has a 64% frequency of reading?

This is the end of the test. Please hand in all your materials. Be sure your name and social security number are on your answer sheet.

(APPENDIX B - TESTBOOK OF PILOT VERSION)

A TEST OF ACADEMIC LISTENING SKILLS

PART 1: Library Tour and Campus Tour

INSTRUCTIONS: You will hear a series of conversations taking place during a library tour and a campus tour. Listen carefully to specific information such as directions, procedures, fares and schedules. You may take notes on the pages provided. You will be asked questions about the information later on.

SITUATION I: Library Tour-conversation between a librarian and students

101. a. show a fee card
b. wait for a month
c. go to a postoffice
d. sign a circulation slip
102. a. new books
b. magazines
c. encyclopedias
d. videotapes
103. a. There is never a problem with renewal requests.
b. The lending library is unlikely to grant a renewal request.
c. Requesting renewal on the due date is discouraged.
d. You will be notified if the renewal is approved.

Now, listen to the next part of the library tour.

104. a. Training faculty in physics.
b. Supporting graduate study in science.
c. Encouraging students to study physics.
d. Keeping a collection of materials.
105. a. Saturday evening
b. Friday evening
c. Sunday morning
d. Sunday evening

Now, listen to the next part of the tour.

106. a. Searchline uses electronic devices to search bibliography lists.
b. You can find research topics quickly.
c. A professional will perform the search for you.
d. Searchline requires extensive knowledge of data processing.
107. a. study a search manual
b. search through related topics
c. select a suitable search area
d. expand topics with specific words
108. a. $(2 \times \$0.50) + (2 \times \$2.50)$
b. $(2 \times \$0.40) + (3 \times \$2.50)$
c. $(3 \times \$0.40) + (3 \times \$2.40)$
d. $(3 \times \$0.50) + (2 \times \$2.50)$

SITUATION II: Campus Tour--conversation between a tour guide and students. As you listen, look at the campus map in your test book. You may write notes on the map. You will be asked questions about it later.

109. a. treasurer's office
b. registrar's office
c. graduate school office
d. admissions office
110. a. 4
b. 7
c. 10
d. 15
111. a. student center
b. field house
c. university dormitory
d. health center
112. a. 11
b. 12
c. 4
d. 5
113. a. 6
b. 9
c. 11
d. 14
114. a. North Grand Hall
b. Public Administration building
c. Field House
d. Health Center

Now, listen as the tour guide tells the students about the Health Center. You may take notes.

- 115. a. surgery
b. health education
c. outpatient clinic
d. diet counseling
- 116. a. \$100
b. \$50
c. \$72
d. \$122
- 117. a. All university students must pay it
b. It covers limited services
c. It is paid once each academic year
d. It is covered by health insurance

Now, listen as the guide tells the students about bus service. You may take notes.

- 118. a. a transfer
b. an unlimited pass
c. a coupon
d. a fee card
- 119. a. They can be requested any time during the trip.
b. They must be used to ride on the first connecting bus.
c. There is a special charge for transfers.
d. They are valid for completing a round trip.
- 120. a. blind people
b. elderly people
c. people accompanying the handicapped
d. children under twelve

PART 2: Student-Teacher Conversations

INSTRUCTIONS: You will hear a series of conversations taking place in a classroom between a teacher and a student. Listen carefully. You will be asked about the meaning of the conversation, or about a proper response. Choose the best answer by marking the letter on your answer sheet. First, listen to an example question:

You will hear...

S: Do we need to memorize this equation?

T: I hope not. I'll talk about it a little more, I want to make sure you understand it, because once you understand you won't need to memorize it.

What is the teacher's intention by giving such a response?

You will read...

- a. giving a command
- b. giving advice
- c. making a request
- d. making a complaint

The best answer to the example is b--giving advice.

Now, listen to the first conversation and answer Question 121 in your test book. The questions are not in the test book so listen carefully.

-
121. a. making a complaint
b. asking for an explanation
c. giving a command
d. offering a suggestion
122. a. I don't know either.
b. So I guess it will be on the test.
c. So we will talk about it later?
d. I sure hope not.

123. a. giving advice
b. giving a command
c. stating a fact
d. showing disappointment
124. a. making an assignment
b. asking for information
c. asking for approval
d. making a proposal
125. a. no interruptions are permitted
b. there is a better alternative
c. asking questions is encouraged
d. there is a disagreement
126. a. his policy is relaxed
b. they should attend every class
c. he has no preference
d. it is all right to be late
127. a. I'll see you at ten.
b. It couldn't be better.
c. When would be a better time?
d. Where should we meet?
128. a. He doesn't give pop quizzes.
b. It's a poor suggestion.
c. The idea needs some consideration.
d. He probably will give pop quizzes.

PART 3: Lectures

INSTRUCTIONS: You will hear some short lectures taking place in a classroom or a lab. Listen carefully to main ideas, supporting ideas, and specific information about processes and functions. You may take notes on the pages provided in your test book. After you have listened to the lectures, you will be asked some questions about them.

Now, listen to lecture I on the Sewage Disposal Unit.

Look at the diagram in your test book. You may write notes on the diagram.

129. a. 1
b. 2
c. 8
d. 10
130. a. 6
b. 7
c. 8
d. 9
131. a. polluted water is pumped through the pump house
b. polluted water is filtered
c. chemicals are added to the water
d. solids settle to the bottom
132. a. soil
b. undecayed solids
c. a filter
d. the settling tank
133. a. to allow the water to evaporate
b. to limit the quantity of water in the filters
c. to keep undecayed solids on the bottom
d. to aerate the water
-

Now listen to lecture II on the Laser Jet Printer. Look at the diagram in your test book. You may write notes on the diagram.

134. a. 5
b. 6
c. 7
d. 8
135. a. 3
b. 4
c. 6
d. 9
136. a. press the on-line key
b. slide the cartridge into the slot
c. set the printer off-line
d. grasp the edge of the cartridge
137. a. plug in the interface cable
b. adjust the print Density Dial
c. connect the AC power cord to a computer
d. insert a font cartridge
138. a. turn #10 clockwise
b. turn #11 counterclockwise
c. turn #10 counterclockwise
d. turn #11 clockwise
139. a. turn it off and begin again
b. consult the manual
c. find someone to repair it
d. open the upper part of the printer

Now, listen to lecture 3 about Technical Writing. Look at the graph in your test book. You may write notes on the graph.

140. a. The body of a report
b. How the readers read
c. Different types of readers
d. Analyzing the reader's needs
141. a. writing a long body
b. disregarding his/her audience
c. not organizing the writing
d. emphasizing one's own field
142. a. one having a concise body
b. one focusing on the summary
c. one providing selective information
d. one taking a short time to read

143. a. summary
b. conclusion
c. introduction
d. body
144. a. appendix
b. body
c. conclusion
d. introduction
145. a. summary
b. appendix
c. introduction
d. conclusion

This is the end of the test. Please hand in all your materials. Be sure your name and social security number are on your answer sheet.

(APPENDIX C - TESTBOOK OF FINAL VERSION)

A TEST OF ACADEMIC LISTENING SKILLS

PART 1: Library Tour and Campus Tour

INSTRUCTIONS: You will hear a series of conversations taking place during a library tour and a campus tour. Listen carefully to specific information such as directions, procedures, fares and schedules. You may take notes on the pages provided. You will be asked questions about the information later on.

SITUATION I: Library Tour-conversation between a librarian and students

101. a. show a fee card
b. wait for a month
c. go to a postoffice
d. sign a circulation slip
102. a. new books
b. periodicals
c. encyclopedias
d. videotapes
103. a. There is never a problem with renewal requests.
b. The lending library is unlikely to grant a renewal request.
c. Requesting renewal on the due date is discouraged.
d. You will be notified if the renewal is approved.

Now, listen to the next part of the library tour.

104. a. electing science materials for students.
b. Supporting graduate study in science.
c. Encouraging students to study physics.
d. Keeping a collection of materials.
105. a. Saturday evening
b. Friday evening
c. Sunday morning
d. Sunday evening

Now, listen to the next part of the tour.

106. a. study a bibliography list
b. search through related topics
c. select a suitable search area
d. expand topics with specific words
107. a. (2 x \$.50) + (2 x \$2.50)
b. (2 x \$.40) + (3 x \$2.50)
c. (3 x \$.40) + (3 x \$2.40)
d. (3 x \$.50) + (2 x \$2.50)

SITUATION II: Campus Tour--conversation between a tour guide and students. Now, listen as the tour guide tells the students about the Health Center. You may take notes.

108. a. surgery
b. health education
c. outpatient clinic
d. diet counseling
109. a. \$100
b. \$50
c. \$72
d. \$122
110. a. All university students must pay it
b. It covers limited services
c. It is paid once each academic year
d. It is covered by health insurance

Now, listen as the guide tells the students about bus service. You may take notes.

111. a. They can be requested any time during the trip.
b. They must be used to ride on the first connecting bus.
c. There is a special charge for transfers.
d. They are valid for completing a round trip.
112. a. blind people
b. elderly people
c. people accompanying the handicapped
d. children under twelve

PART 2: Student-Teacher Conversations

INSTRUCTIONS: You will hear a series of conversations taking place in a classroom between a teacher and a student. Listen carefully. You will be asked about the meaning of the conversation, or about a proper response. Choose the best answer by marking the letter on your answer sheet. First, listen to an example question:

You will hear...

S: Do we need to memorize this equation?

T: I hope not. I'll talk about it a little more, I want to make sure you understand it, because once you understand you won't need to memorize it.

What is the teacher's intention by giving such a response?

You will read...

- a. giving a command
- b. giving advice
- c. making a request
- d. making a complaint

The best answer to the example is b—giving advice.

Now, listen to the first conversation and answer Question 113 in your test book. The questions are not in the test book so listen carefully.

- 113.
 - a. making a complaint
 - b. asking for an explanation
 - c. giving a command
 - d. offering a suggestion
- 114.
 - a. I don't know either.
 - b. So I guess it will be on the test.
 - c. So we will talk about it later?
 - d. I sure hope not.

115. a. giving advice
b. giving a command
c. stating a fact
d. showing disappointment
116. a. making an assignment
b. asking for information
c. asking for approval
d. making a proposal
117. a. no interruptions are permitted
b. there is a better alternative
c. asking questions is encouraged
d. there is a disagreement
118. a. his policy is fair
b. they should attend every class
c. he has no preference
d. it is all right to be absent
119. a. I'll see you at ten.
b. It couldn't be better.
c. When would be a better time?
d. Really, I hope you can make it.
120. a. He doesn't give pop quizzes.
b. It's a poor suggestion.
c. The idea needs some consideration.
d. He probably will give pop quizzes.

PART 3: Lectures

INSTRUCTIONS: You will hear some short lectures taking place in a classroom or a lab. Listen carefully to main ideas, supporting ideas, and specific information about processes and functions. You may take notes on the pages provided in your test book. After you have listened to the lectures, you will be asked some questions about them.

Now, listen to Lecture I on the Sewage Disposal Unit.

Look at the diagram in your test book. You may write notes on the diagram.

- 121. a. 1
b. 2
c. 8
d. 10
 - 122. a. 6
b. 7
c. 8
d. 9
 - 123. a. polluted water is pumped through the pump house
b. polluted water is filtered
c. chemicals are added to the water
d. solids settle to the bottom
 - 124. a. soil
b. undecayed solids
c. a filter
d. the settling tank
 - 125. a. to allow the water to evaporate
b. to limit the quantity of water in the filters
c. to keep undecayed solids on the bottom
d. to aerate the water
-

Now listen to Lecture II on the Laser Jet Printer. Look at the diagram in your test book. You may write notes on the diagram.

- 126. a. 5
b. 6
c. 7
d. 8
- 127. a. 3
b. 4
c. 6
d. 9
- 128. a. press the on-line key
b. slide the cartridge into the slot
c. set the printer off-line
d. grasp the edge of the cartridge
- 129. a. plug in the interface cable
b. adjust the print Density Dial
c. connect the AC power cord to a computer
d. insert a font cartridge
- 130. a. turn it off and begin again
b. consult the manual
c. find someone to repair it
d. open the upper part of the printer

Now, listen to Lecture III about Technical Writing. Look at the graph in your test book. You may write notes on the graph.

- 131. a. The body of a report
b. Technical report writers
c. Different types of readers
d. Analyzing the reader's needs
- 132. a. writing a long body
b. disregarding his/her audience
c. not organizing the writing
d. emphasizing one's own field
- 133. a. summary
b. conclusion
c. introduction
d. body
- 134. a. appendix
b. body
c. conclusion
d. introduction

135. a. summary
b. appendix
c. introduction
d. conclusion

This is the end of the test. Please hand in all your materials. Be sure your name and social security number are on your answer sheet.

(APPENDIX D - STEPS OF ADMINISTERING THE NEW ESLPT LISTENING COMPREHENSION TEST)

1. Inform the students that they will be given a test of academic listening skills. The test will take about 30 minutes.
2. Inform the students the test is tape-recorded, it is divided into 3 parts, and there are 35 multiple-choice format test items.
3. Inform the students that they ought to put down the answers on the same answer sheet that they have been using, starting from number "101".
4. Inform the students that they will be given 2 testbooks. one with the answers and another with a notetaking page and 3 diagrams, one for each lecture in part 3. The diagrams are in the same order as the lectures on the tape.
5. Inform the students that they can write notes on the notetaking page, and on the diagrams too.
6. Pass out the testbook with answers first, then, the one with the diagrams.
7. Ask if every student has been given both testbooks, and ask them to make sure that two testbook numbers are matched.
8. Inform the students that you are going to play the tape in the next minute, and you will not stop the tape during the test.
9. Test out the tape, make sure that students can hear it pretty well every where. Adjust the volume when needed.
Ask if there are any questions.
10. Play the tape through with no stops.
11. After the test is over, collect all three things from the students: the testbook with answers, the testbook with diagrams and a notetaking page, and the computer-scoring

answer sheet.

12. Wait till every student has turned in all the materials before dismissing them.